PATENT ABSTRACTS OF JAPAN

(11)Publication number : **2003-058196**
(43)Date of publication of application : **28.02.2003**

(51)Int.Cl. G10L 19/02
G10L 19/00
H03M 7/30

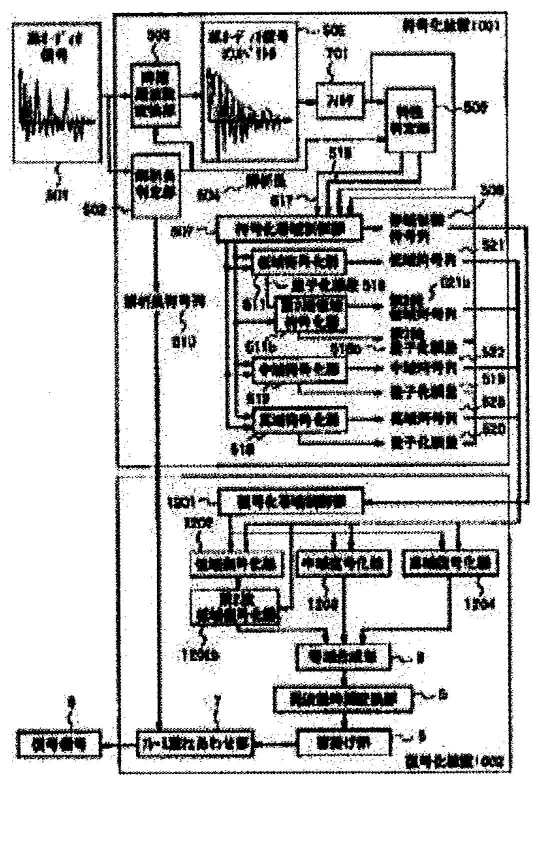
(21)Application number : **2002-211570**
(22)Date of filing : **11.03.1999**

(71)Applicant : **MATSUSHITA ELECTRIC IND CO LTD**
(72)Inventor : **TSUSHIMA MINEO**
NORIMATSU TAKESHI
ISHIKAWA TOMOKAZU

(30)Priority
Priority number : **10059403** Priority date : **11.03.1998** Priority country : **JP**

(54) AUDIO SIGNAL ENCODING METHOD AND AUDIO SIGNAL DECODING METHOD

(57)Abstract:
PROBLEM TO BE SOLVED: To provide an audio signal encoding method and an audio signal decoding method for adaptively performing scalable coding so that sufficient performance can be presented when encoding various kinds of audio signals.
SOLUTION: Adaptive scalable coding is performed for changing the frequency range of encoding by automatically deciding the character/distribution of source audio signals without performing fixed scalable coding.



CLAIMS

[Claim(s)]
[Claim 1]A characteristic determination step, a coding bandwidth control step, and a coding step are included, Time - Are an audio signal by which frequency conversion was carried out an audio signal encoding method changed into a coding row, and a coding row, Including encoded information and a bandwidth control code sequence, a coding step, Have two or more coding sub steps, control of a coding bandwidth control step performs multistage coding of an audio signal, output encoded information, and a characteristic determination step, Output zone weight information which judges an audio signal inputted and shows a weighting of each frequency band to code, and a coding bandwidth control step, A quantization zone of each coding sub step which constitutes multistage coding based on zone weight information, An audio signal encoding method which determines connection order, makes multistage coding constituted scalable based on a quantization zone of each determined coding sub step, and connection order perform to a coding step, and outputs a bandwidth control code sequence which shows a quantization zone of each determined coding sub step, and connection order.
[Claim 2]The audio signal encoding method according to claim 1 which determines a quantization zone of each coding sub step, and connection order so that a coding bandwidth control step may become either of the multistage coding defined beforehand.
[Claim 3]The audio signal encoding method according to claim 1 with which a coding step outputs a quantization error and a coding bandwidth control step determines a quantization zone of each coding sub step, and connection order based on zone weight information and a quantization error.
[Claim 4]It is an audio signal decoding method which decodes a coding row which includes a decoding bandwidth control step and a decoding step, and includes encoded information and a bandwidth control code sequence to an audio signal, A bandwidth control code sequence is shown and a quantization zone of each coding at the time of carrying out the multistage coding of the encoded information, and connection order a decoding step, An audio signal decoding method which has two or more decoding sub steps, and performs a multistage decoding of encoded information by control of a decoding bandwidth control step and with which a decoding bandwidth control step makes a multistage decoding constituted scalable based on a bandwidth control code sequence perform to a decoding step.

[Translation done.]

DETAILED DESCRIPTION

[Detailed Description of the Invention]
[0001]
[Field of the Invention]This invention about an audio signal encoding method and an audio signal decoding method, The changed signal is compared with an original audio signal using the characteristic quantity especially obtained from audio signals, such as an audio signal and a music signal, especially the signal which changed the audio signal into the frequency domain from the segment of time using techniques, such as an orthogonal transformation, High quality and a broadband audio signal are related with the decoding method of the composition which can be decoded only using the methods of coding efficiently, in order to express by the fewest possible code sequences, all the coding rows that are the coded signals, or its part.
[0002]
[Description of the Prior Art]The various techniques of coding and decrypting an audio signal efficiently are proposed. In the compression encoding type of the audio signal which has not less than 20-kHz frequency bands, such as a music signal, it is [an MPEG audio system and] Twin VQ. There are a system (TC-WVQ) etc. The coding mode represented by the MPEG system is a system which changes the digital audio signals of a time-axis into the data on a frequency axis using orthogonal transformations, such as cosine transformation, and codes the information on the frequency axis from important information auditorily using human being's auditory sensitivity characteristic.
Unimportant information and redundant information are systems which are not coded auditorily.
On the other hand, it is Twin VQ. A system (TC-WVQ) has a coding mode which it is going to express for the quite small amount of information to the amount of information of an original digital signal using the vector quantization technique. An MPEG audio and Twin VQ (TC-WVQ), Respectively ISO/IEC standard IS-11172-3 and T.Moriya, H.Suga:An 8 Kbits transform coder for noisy channels, and Proc.ICASSP. It is stated to 89, pp196-199, etc.
[0003]Here, the general outline of a Twin VQ system is explained using drawing 10. The original audio signal 101 is inputted into the analysis length judgment part 102, and analysis length is computed. Simultaneously, the analysis length judgment part 102 quantizes the analysis length 112, and outputs the analysis length code sequence 111. Next, according to the analysis length 112, the original audio signal 101 is changed into the original audio signal 104 of a frequency domain by the temporal modulation converter 103. Next, the normalizing process (flattening processing) of the original audio signal 104 of a frequency domain is carried out by the normalization processing part (flattening treating part) 106, and it obtains the audio signal 108 after a normalizing process. A normalizing process calculates the frequency facies 105 from the original audio signal 104, and is performed by breaking by the frequency facies 105 which computed the original audio signal 104. The normalization processing part 106 quantizes the frequency facies information that it used for the normalizing process, and outputs the normalization code sequence 107. Next, the audio signal 108 after a normalizing process is quantized by the vector quantization part 109,

and the code sequence 110 is acquired.

[0004]Even if it uses a part of code sequence inputted into a decoder in recent years, there is a thing with the structure which can reproduce an audio signal. The above-mentioned structure is called scalable structure and it calls it scalable coding to code so that scalable structure can be realized.

[0005]An example of fixed scalable coding adopted as drawing 11 by the general Twin VQ system is shown. According to the analysis length 1314 judged by the analysis length judgment part 1303, the original audio signal 1304 of a frequency domain is obtained from the original audio signal 1301 by the temporal modulation converter 1302. Next, if the original audio signal 1304 of a frequency domain is inputted into the low-pass coding equipment 1305, the quantization error 1306 and the low-pass code sequence 1311 will be outputted. If the quantization error 1306 is inputted into the mid-range coding equipment 1307, the quantization error 1308 and the mid-range code sequence 1312 will be outputted. If the quantization error 1308 is inputted into the high region coding equipment 1309, the quantization error 1310 and the high region code sequence 1313 will be outputted. Here the above-mentioned low-pass one, a mid-range, or high region coding equipment, Having a normalization processing part and a vector quantization part, the output outputs low-pass and a mid-range including each code sequence outputted by the quantization error, the normalization processing part, and the vector quantization part, or a high region code sequence.

[0006]

[Problem to be solved by the invention]Since low-pass, a mid-range, and each zone quantizer of the high region were being fixed in fixed scalable coding of the conventional system as shown in drawing 11, as shown in drawing 12, it was difficult to code to distribution of an original audio signal, so that a quantization error may be lessened as much as possible. So, when coding an audio signal with various character and distribution, it was difficult to be unable to demonstrate sufficient performance but to perform efficient scalable coding by high-quality sound.

[0007]When it was made in order that this invention might cancel the above-mentioned problem, and coding an audio signal when coding various audio signals, as shown in drawing 13, It aims at providing the audio signal encoding method which it is efficient, and is a low bit rate, and can code to high-quality sound, and an audio signal decoding method by carrying out scalable coding of the various audio signals accommodative.

[0008]

[Means for solving problem]In order to solve this SUBJECT, the audio signal encoding method concerning this invention, and an audio signal decoding method, Not using fixed scalable coding, it is made to perform adaptation scalable coding to which the character of an original audio signal and the frequency range coded in accordance with distribution are changed.

[0009]The audio signal encoding method concerning this invention A characteristic determination step, a coding bandwidth control step, A coding step is included and it is time. - Are the audio signal by which frequency conversion was carried out an audio signal encoding method changed into a coding row, and a coding row, Including encoded information and a bandwidth control code sequence, a coding step, Have two or more coding sub steps, control of a coding bandwidth control step performs multistage coding of an audio signal, output encoded information, and a characteristic determination step, Output the zone weight information which judges the audio signal inputted and shows the weighting of each frequency band to code, and a coding bandwidth control step, The quantization zone of each coding sub step which constitutes multistage coding based on zone weight information, Connection order is determined, the multistage coding constituted scalable based on the quantization zone of each determined coding sub step and connection order is made to perform to a coding step, and the bandwidth control code sequence which shows the quantization zone of each determined coding sub step and connection order is outputted.

[0010]In said audio signal encoding method, an audio signal encoding method concerning this invention determines a quantization zone of each coding sub step, and connection order so that a coding bandwidth

control step may become either of the multistage coding defined beforehand.

[0011]In said audio signal encoding method an audio signal encoding method concerning this invention, A coding step outputs a quantization error and a coding bandwidth control step determines a quantization zone of each coding sub step, and connection order based on zone weight information and a quantization error.

[0012]An audio signal decoding method concerning this invention A decoding bandwidth control step, Are a coding row which includes a decoding step and includes encoded information and a bandwidth control code sequence an audio signal decoding method decoded to an audio signal, and a bandwidth control code sequence, A quantization zone of each coding at the time of carrying out the multistage coding of the encoded information and connection order are shown, and a decoding step, It has two or more decoding sub steps, control of a decoding bandwidth control step performs a multistage decoding of encoded information, and it is made for a decoding bandwidth control step to make a multistage decoding constituted scalable based on a bandwidth control code sequence perform to a decoding step.

[0013]

[Mode for carrying out the invention]Hereafter, the embodiment of the invention 1 is described about Embodiment 2 using drawing 14 thru/or drawing 20, using drawing 1 thru/or drawing 9.

[0014](Embodiment 1) Drawing 1 shows a block diagram of an audio signal encoding device by the embodiment of the invention 1 which performs adaptation scalable coding. In drawing 1, 1001 is coding equipment which codes the original audio signal 501. An analysis length judgment part which judges the analysis length 504 at the time of 502 analyzing the above-mentioned field audio signal 501 in this coding equipment 1001, and 503 are the units of the above-mentioned analysis length 504, A temporal modulation converter which changes a time-axis of the original audio signal 501 into a frequency axis, Analysis length by which 504 was judged by the above-mentioned analysis length judgment part 502, and 505 A spectrum of an original audio signal, A filter into which the spectrum 505 of this original audio signal is inputted 701, Each coding equipment 511,512,513,511b of two or more stages of each which 506 judges the characteristic of the spectrum 505 of an original audio signal, and can be set to the above-mentioned coding equipment 1001, A characteristic judgment part which determines a frequency band of an audio signal which ********, and a frequency band of each coding equipment in which 507 was determined by this characteristic judgment part 506, Consider an audio signal by which frequency conversion was carried out [above-mentioned] as the input, and Each coding equipment 512,513,514,511b of two or more stages of each, A bandwidth control code sequence which is the above-mentioned code sequence which determines ********, and to which a quantization zone of each coding equipment and a coding band control part which changes connection order into a code sequence, and 508 are outputted from this coding band control part 507, An analysis length code sequence made into a code sequence and 511,512,513 the above-mentioned analysis length 504 to which 510 was outputted from the above-mentioned analysis length judgment part 502, Low-pass coding equipment which was mentioned above and which codes low-pass, a mid-range, and a signal of a high region, respectively, Mid-range coding equipment, high region coding equipment, the 2nd step low-pass coding equipment in which 511b codes the quantization error 518 of the low-pass coding equipment 511 of the 1st step, A low-pass code sequence which is a coded signal with which 521,522,523 is outputted from this each coding equipment 511,512,513, A mid-range code sequence, a high region code sequence, the 2nd step low-pass code sequence whose 521b is an encoding output of the 2nd step low-pass coding equipment 511b, A quantization error which is a difference of a signal before being coded and each above-mentioned coded signal with which 518,519,520 is outputted from this each coding equipment 511,512,513, and 518b are the 2nd step quantization errors which are quantization errors of the 2nd step low-pass coding equipment 511b.

[0015]On the other hand, 1002 is a decoding device which decrypts the coding row coded with the above-mentioned coding equipment 1001. The frequency time converter which performs conversion with 5 [contrary to the temporal modulation converter 503 in the above-mentioned coding equipment 1001] in this

decoding device 1002, The window credit part which performs window credit which 6 multiplies by a windowing function on a time-axis, and 7 A frame superposition part, 8 a decoded signal and 9 a band composition part and 1201 a decoding band control part, and 1202-1203 and 1204, The low-pass decoding machine which decrypts corresponding to the above-mentioned low-pass coding equipment, mid-range coding equipment, and the high region coding equipment 511,512,513, respectively, a mid-range decoding machine, a high region decoding machine, and 1202b are the 2nd step low-pass decoding machines which decrypt the output of the 1st step low-pass decoding machine 1202.
[0016]here -- the zone of further others [machine / the coding equipment after the 2nd step, and / decoding] -- the accuracy of coding and a decoding can be improved if needed, so that it may provide also in multistage and this becomes multistage.
[0017]Hereafter, operation of the coding equipment 1001 is explained first. The original audio signal 501 which it is going to code presupposes that it is a digital signal series which continues in time. For example, suppose that it is an audio signal the digital signal quantized to 16 bits by 48 kHz of sampling frequencies.
[0018]The above-mentioned field audio signal 501 is inputted into the analysis length judgment part 502. The above-mentioned analysis length judgment part 502 judges the characteristic of the inputted above-mentioned field audio signal 501, it opts for the analysis length 504, and the result is sent to the decoding device 1002 as the analysis length code sequence 510. As the analysis length 504, 256, 1024, 4096, etc. are used, for example. For example, when the high region frequency component contained in the original audio signal 501 exceeds a predetermined value. The analysis length 504 is set to 256, a low-pass frequency component exceeds a predetermined value, and when a high region frequency component is smaller than a predetermined value, the analysis length 504 is set to 4096, and when other, the analysis length 504 is set to 1024. In this way, according to the determined analysis length 504, the spectrum 505 of the original audio signal 501 is computed by the temporal modulation converter 503.
[0019]The block diagram of the temporal modulation converter 503 in the audio signal encoding device by the embodiment of the invention 1 is shown in drawing 2. The above-mentioned field audio signal 501 will output, if it is accumulated by the frame dividing part 201 and the this accumulated measurement size reaches the analysis length 504 determined by the above-mentioned analysis length judgment part 502 until the sampled value reaches a predetermined measurement size. In the case where the frame dividing part 201 is a thing of composition of outputting for every shift length of a certain, for example, the analysis length 504 is made into 4096 samples, If the shift length of the half of the analysis length 504 is set up, the analysis length 504 has the composition of outputting the 4096 newest samples for every time to be equivalent to reaching 2048 samples. Though natural, even if the analysis length 504 and a sampling frequency change, it is possible to have similarly the composition which set shift length as the half of the analysis length 504. And the output from this frame dividing part 201 is inputted into the latter window credit part 202. In the window credit part 202, to the output from the frame dividing part 201, it multiplies by a windowing function on a time-axis, and is considered as the output of the window credit part 202. This situation is shown by (several 1), for example.
[0020]
[Mathematical formula 1]

$$hxi = h_i * x_i \quad i = 1,2,\cdots,N$$

$$h_i = \sin\left\{\frac{\pi}{N}(i+0.5)\right\}$$

However, xi is an output from the frame dividing part 201, hi is a windowing function and hxi is an output from the window credit part 202 here. i is still a suffix of time. The windowing function hi shown by (several 1) needs to be an example, and the windowing function does not necessarily need to be a thing of (several 1).

[0021]It depends for selection of a windowing function on the feature of the signal inputted into the window credit part 202, the analysis length 504 of the frame dividing part 201, and the form of the windowing function in the frame located forward and backward in time. For example, as a feature of the signal inputted into the window credit part 202, when the analysis length 504 of the frame dividing part 201 is set to N, the average power of the signal inputted for every N/4 is computed, When changing the average power very sharply, it chooses performing the operation shown for making the analysis length 504 shorter than N (several 1) etc. Choosing suitably is desirable so that there may be no distortion in the form of the windowing function of the frame of the present time according to the form of the windowing function of the frame of the last time, and the form of the windowing function of a back frame.
[0022]Subsequently, the output from the window credit part 202 is inputted into the MDCT section 203, a modification discrete cosine transform is given here, and a MDCT coefficient is outputted. The general formula of a modification discrete cosine transform is expressed with (several 2).
[0023]
[Mathematical formula 2]

$$y_k = \sum_{n=0}^{N-1} hx_n * \cos\left\{\frac{2\pi\left(k+\frac{1}{2}\right)(n+n_0)}{N}\right\}$$

$$n_0 = \frac{N}{4} + \frac{1}{2} \quad \left(k = 0, 1, \cdots, \frac{N}{2} - 1\right)$$

Thus, the MDCT coefficient which is an output of the MDCT section 203, (Several 2) Supposing it can express with inner yk, the output of the MDCT section 203 shows a frequency characteristic, and it corresponds to a high frequency component at linearity, so that it increases from 0 to a low frequency component, so that its variable k of yk is [zero] near, and it becomes close to N/2-1. In this way, the computed above-mentioned MDCT coefficient serves as the spectrum 505 of an original audio signal.
[0024]Next, the spectrum 505 of the above-mentioned field audio signal is inputted into the filter 701. If this filter 701 is inputted as x701(i) and an output is made into y701(i), the filter expressed with (several 3) will be used, for example.
[0025]
[Mathematical formula 3]

$$y_{701}(i) = w_{701}(i) * \{x_{701}(i) + x_{701}(i+1)\}$$
$$i = 0, 1, \cdots, fs-2$$

Here, fs is the analysis length 504. Although the filter 701 expressed with (several 3) is a kind of moving average filter, though natural, it is not necessary to limit it to a moving average filter, they may be other high pass filters, and may be a zone control filter.
[0026]The output of the filter 701 and the analysis length 504 which computed by the analysis length judgment part 502 are inputted into the characteristic judgment part 506. The details of the characteristic judgment part 506 are shown in drawing 6. In the characteristic judgment part 506, the auditory and physical characteristic of spectrum 505** of the original audio signal 501 and an original audio signal is determined. Auditory and the physical characteristic of the original audio signal 501 and this spectrum 505 are the differences between a sound, music, or **, for example. In the case of a sound, most frequency components are, for example on low-pass from 6 kHz.
[0027]Next, operation of the characteristic judgment part 506 is explained using drawing 6. If a signal which filtered the spectrum 505 of an original audio signal inputted into the characteristic judgment part 506 with the filter 701 is made into x506(i), Based on this x506(i), spectrum power p506(i) is calculated by the spectrum power calculation part 803 by (several 4).

[0028]
[Mathematical formula 4]
$p_{506}(i) = x_{506}(i)^2$

Set this spectrum power p506(i) to one of the inputs of the coding band control part 507, and let it be the bandwidth control dignity 517 of each coding equipment. When the analysis length 504 is small, for example it becomes 256, it determines by the arrangement deciding part 804, and the coding zone arrangement information 516 is sent to the coding band control part 507 as fixed location so that each coding equipment may be arranged fixed.
[0029]When the analysis length 504 is small and it is except, it determines by the arrangement deciding part 804 at the time of 4096 or 1024, and it sends the coding zone arrangement information 516 to the coding band control part 507 as dynamic allocation, for example so that each coding equipment may be arranged dynamically.
[0030]Next, operation of the coding band control part 507 is explained using drawing 7. The bandwidth control dignity 517 which is an output from the above-mentioned characteristic judgment part 506 at the coding band control part 507, A signal which filtered the coding zone arrangement information 516 and the spectrum 505 of an original audio signal with the filter 701, the quantization error 518 which each coding equipment outputted, or 519 or 520 is inputted. However, there are these inputs, in order that each coding equipment 511, 512, 513, and 511b and the coding band control part 507 may operate recursively, and in operation of the first-time coding band control part 507, since there is no quantization error, it becomes three inputs except a quantization error.
[0031]As mentioned above, when the analysis length 504 is small and the coding zone arrangement information 516 is placed in a fixed position, According to fixed location of a zone defined beforehand, it codes by determining a quantization zone of coding equipment, the number, and connection order by the quantization sequence decision section 902 and the number deciding part 903 of coding equipment, and the bandwidth calculation part 901 so that coding may be performed from low-pass in order to a mid-range and a high region. That is, band information, the number of coding equipment, and its connection order foreword of coding equipment are coded as information by the bandwidth control code sequence 508 at that time.
[0032]for example, the coding zone and the number of coding equipment of each coding equipment -- respectively -- it codes by arranging coding equipment to one and 0 Hz - 8 kHz, at 0 Hz - 4 kHz so that it may be set to one and 8 kHz - 16 kHz with two and may become three at one and 4 kHz - 12 kHz at 16 kHz - 24 kHz.
[0033]Next, operation of the coding band control part 507 when the coding zone arrangement information 516 is dynamic allocation is explained. The coding band control part 507 consists of the bandwidth calculation part 901 which determines the quantization bandwidth of each coding equipment, the quantization sequence decision section 902 which determines a quantization order of each coding equipment, and three of number deciding part of coding equipment 903** which determines the number of the coding equipment of each zone further. Although it is a translation which determines the bandwidth of each coding equipment based on the signal inputted into the coding band control part 507, In each zone (a predetermined zone, for example, 0 Hz - 4 kHz, 0 kHz - 8 kHz, 4 kHz - 12 kHz, 8 kHz - 16 kHz, and 16 kHz - 24 kHz), the bandwidth control dignity 517 and the average value of what carried out the multiplication of the quantization error after each coding equipment codes are computed. Here, if bandwidth control dignity 517 is made into weight517(i) and a quantization error is made into err507(i), average value will be computed by (several 5).
[0034]
[Mathematical formula 5]

$$Ave_{901}(j) = \frac{1}{f_{upper}(j) - f_{lower}(j)} \sum_{i=f_{upper}(j)}^{f_{lower}(j)} weight_{517}(i) * err_{507}(i)^2$$

Average value [in / j / here / in an index of each zone, and Ave901 (j) / the zone j], and fupper (j) And flower(s) (j) are upper limited frequency of the zone j, and a lower cut off frequency. In this way, average value Ave901 (j) obtained search j used as the maximum, and it serves as a zone which coding equipment codes. Send a value of searched j to the number deciding part 903 of coding equipment, and the one number of coding equipment of a zone corresponding to j is increased, It memorizes how many coding equipment exist in a predetermined coding zone, and coding is repeated until the sum total of the memorized number of coding equipment becomes a total of coding equipment determined beforehand. Finally, a zone and the number of coding equipment of coding equipment are transmitted to a decoding machine as the bandwidth control code sequence 508.

[0035]Next, operation of the coding equipment 3 is explained using drawing 3. The coding equipment 3 consists of the normalizing part 301 and the quantizing part 302. In the normalizing part 301, a MDCT coefficient is normalized using some parameters by considering both with a MDCT coefficient which are a signal of a time-axis which is an output from the frame dividing part 201, and an output from the MDCT section 203 as an input. Normalization of a MDCT coefficient means here oppressing dispersion in a size of a MDCT coefficient which has a difference in a size dramatically by low-pass ingredient and a high-frequency component, For example, to a high-frequency component, a low-pass ingredient points out oppressing dispersion in a size of a MDCT coefficient of a low-pass ingredient by electing a parameter which serves as a big value and a value small in a high-frequency component, and doing division of the above-mentioned MDCT coefficient now, when very large. In the normalizing part 301, an index expressing a parameter used for normalization is coded as the normalization code sequence 303.

[0036]In the quantizing part 302, the MDCT quantization of coefficient is performed by considering the MDCT coefficient normalized by the normalizing part 301 as an input. Under the present circumstances, this quantizing part 302 outputs this such code IDDEKKUSU to which the difference between the this quantized value and each quantization output corresponding to two or more code indices in a code book becomes the smallest. In this case, the difference of the value quantized by the above-mentioned quantizing part 302 and the value corresponding to the code index outputted from this quantizing part 302 is a quantization error.

[0037]Next, a detailed example of the above-mentioned normalizing part 301 is explained using drawing 4. In drawing 4, the frequency facies normalizing part in which 401 receives the output of the frame dividing part 201 and the MDCT section 203, and 402 are zone amplitude normalizing parts which undergo the output of the above-mentioned frequency facies normalizing part 401, and normalize with reference to the band table 403.

[0038]Next, operation is explained. In the frequency facies normalizing part 401, using the data output on the time-axis from the frame dividing part 201, the frequency facies which are facies of rough frequency are computed, and division of the MDCT coefficient which is an output from the MDCT section 203 is done. The parameter used for expressing frequency facies is coded as the normalization code sequence 303. In the zone amplitude normalizing part 402, it normalizes for every zone shown with the band table 403 by considering the output signal from the frequency facies normalizing part 401 as an input. for example, -- supposing the MDCT coefficient which is an output of the frequency facies normalizing part 401 considers it as dct (i) (i = 0-2047) and the band table 403 is a thing as shown in (Table 1), for example -- etc. (several 6) etc. -- the average value of the amplitude for every zone is computed by using.

[0039]

[Table 1]

帯域 k	$f_{lower}(k)$	$f_{upper}(k)$
0	0	10
1	11	22
2	23	33
3	34	45
4	46	56
5	57	68
6	69	80
7	81	92
8	93	104
9	105	116
10	117	128
11	129	141
12	142	153
13	154	166
14	167	179
15	180	192
16	193	205
17	206	219
18	220	233
19	234	247
20	248	261
21	262	276
22	277	291
23	292	307
24	308	323
25	324	339
26	340	356
27	357	374
28	375	392
29	393	410
30	411	430
31	431	450
32	451	470
33	471	492
34	493	515
35	516	538
36	539	563
37	564	587
38	589	615
39	616	643
40	645	673
41	674	705
42	706	737
43	738	772
44	773	809
45	810	848
46	849	889
47	890	932
48	933	978
49	979	1027
50	1028	1079
51	1080	1135
52	1136	1193
53	1194	1255
54	1256	1320
55	1321	1389
56	1390	1462
57	1463	1538
58	1539	1617
59	1618	1699
60	1700	1783
61	1784	1870
62	1871	1958
63	1959	2048

[Mathematical formula 6]

$$sum_j = \sum_{i=bjlow}^{bjhigh} dct(i)^p$$

$$ave_j = \left(\frac{sum_j}{bjhigh - bjlow + 1} \right)^{-p} \quad bjlow \leq i \leq bjhigh$$

Here, bjlow and bjhigh show most the low-pass index i and the index i of most a high region whose dct(i) in the j-th zone shown in the band table 403 belongs, respectively. p is the norm in distance calculation and 2 etc. is desirable. avej is the average value of amplitude in each zone number j. In the zone amplitude normalizing part 402, qavej is computed by quantizing avej, for example, it normalizes using (several 7).

[0040]
[Mathematical formula 7]

$$n\_dct(i) = dct(i) / gave_j \qquad bjlow \leq i \leq bjhigh$$

Quantization of avej may use quantization of a scalar and may perform vector quantization using a code book. In the zone amplitude normalizing part 402, an index of a parameter used for expressing qavej is coded as the normalization code sequence 303.

[0041]Although composition of the normalizing part 301 in coding equipment showed a thing of composition of having used both the frequency facies normalizing part 401 of drawing 4 and the zone amplitude normalizing part 402, composition only using the frequency facies normalizing part 401 may be used for it, and composition only using the zone amplitude normalizing part 402 may be used for it. It is good also as composition which inputs an output signal of the MDCT section 203 into the quantizing part 302 as it is with composition which does not use above-mentioned both by low-pass ingredient of a MDCT coefficient outputted from the MDCT section 203, and a high-frequency component when there is no big dispersion.

[0042]Next, details of the frequency facies normalizing part 401 of drawing 4 are explained using drawing 5. In drawing 5, a linear-predictive-coding part in which 601 receives an output of the frame dividing part 201, a facies quantizing part in which 602 receives an output of the linear-predictive-coding part 601, and 603 are envelopment characteristic normalizing parts which undergo an output of the MDCT section 203.

[0043]Next, operation of the above-mentioned frequency facies normalizing part 401 is explained with reference to drawing 5. In the above-mentioned linear-predictive-coding part 601, linear predictive coding (Linear Predictive Coding) is conducted by considering an audio signal on a time-axis from the frame dividing part 201 as an input. Linear predictor coefficients (LPC coefficient) of linear predictive coding compute an autocorrelation function of signals by which window credit was carried out, such as a humming window, are solving a normal equation etc. and, generally can compute it. Computed linear predictor coefficients are changed into a line spectrum pair coefficient (LSP (LineSpectrum Pair) a coefficient) etc., and are quantized by the facies quantizing part 602. As the quantization technique here, vector quantization may be used and scalar quantity child-ization may be used. And a frequency transfer characteristic which a parameter quantized by the facies quantizing part 602 expresses is computed by the envelopment characteristic normalizing part 603, and a MDCT coefficient which is an output from the MDCT section 203 is normalized by doing division now. If linear predictor coefficients equivalent to a parameter quantized by the facies quantizing part 602 as a concrete example of calculation are made into qlpc(i), the above-mentioned frequency transfer characteristic computed by the envelopment characteristic normalizing part 603 can be expressed with (several 8), for example.

[0044]
[Mathematical formula 8]

$$li = \begin{cases} qlpc(i) & 0 \leq i \leq ORDER \\ 0 & ORDER+1 \leq i \leq N \end{cases}$$

$$env(i) = \frac{1}{fft(li)}$$

Here, as for ORDER, ten to about 40 are desirable. fft() means Fast Fourier Transform. using computed frequency transfer characteristic env(i) -- the envelopment characteristic normalizing part 603 -- the following -- being shown (several 9) -- it normalizes by using.

[Mathematical formula 9]

$$fdct(i) = \frac{mdct(i)}{env(i)}$$

Here, mdct(i) is an output signal from the MDCT section 203, and fdct(i) is an output signal from the normalized envelopment characteristic normalizing part 603.
[0045]Next, detailed operation of the quantization method of the quantizing part 302 in the above-mentioned coding equipment 1 is explained using drawing 8. The MDCT coefficient 1001 inputted into the quantizing part 302 extracts some from the MDCT coefficient 1001, and constitutes the sound-source subvector 1003. Similarly when the coefficient sequence which divided the MDCT coefficient which is an input of the normalizing part 301 by the MDCT coefficient which is an output of the normalizing part 301 is used as the normalization ingredient 1002, by the normalizing part 301 also about this normalization ingredient 1002. A subvector can be extracted from this normalization ingredient 1002, and the dignity subvector 1004 can consist of same rules as having extracted the sound-source subvector 1003 from the MDCT coefficient 1001. The rule which extracts the sound-source subvector 1003 and the dignity subvector 1004 from the MDCT coefficient 1001 and the normalization ingredient 1002, respectively has a method etc. which are shown by (several 10), for example.
[0046]
[Mathematical formula 10]

$$\mathrm{subvector}_i(j) = \begin{cases} \mathrm{vector}\left(\frac{\mathrm{VTOTAL}}{\mathrm{CR}} \cdot i + j\right) \text{ ただし } \frac{\mathrm{VTOTAL}}{\mathrm{CR}} * i + j < \mathrm{TOTAL} \\ 0 \quad \text{ただし } \frac{\mathrm{VTOTAL}}{\mathrm{CR}} * i + j \geq \mathrm{TOTAL} \end{cases}$$

The j-th element of the i-th sound-source subvector is subvector i (j), the MDCT coefficient 1001 is vector() here, and the total number of elements of the MDCT coefficient 1001 is TOTAL, The number of elements of the sound-source subvector 1003 is the value as TOTAL with same CR and VTOTAL, or a larger value, and it sets up so that VTOTAL/CR may become a positive number value. For example, when TOTAL is 2048, CR is 19, VTOTAL is 2052, CR is 23, VTOTAL is 2070, CR is 21, and VTOTAL is 2079 etc. The dignity subvector 1004 can also be extracted in several 10 procedure. In the vector quantizer 1005, out of the code vector in the code book 1009. The index of a code vector with which the distance with the sound-source subvector 1003 looked for that to which weighting **** also becomes small by the dignity subvector 1004, and gave the minimum distance, The remainder subvector 1010 equivalent to the quantization error of the code vector and the input sound source subvector 1003 which gave the minimum distance is outputted.
[0047]In a actual example of a computational procedure, the vector quantizer 1005 explains as what consists of three components, the distance calculation means 1006, the code determination means 1007, and remainder creating means 1008**. In the distance calculation means 1006, distance of the i-th sound-source subvector 1003 and the k-th code vector of the code book 1009 is computed, for example using (several 11).
[0048]
[Mathematical formula 11]

$$\mathrm{dik} = \sum_{j=0}^{\mathrm{CR}-1} w_j{}^R (\mathrm{subvector}_i(j) - C_k(j))^S$$

Here, the j-th element of the k-th code vector, R, and S are the norm of distance calculation, and, as for the j-th element of a dignity subvector, and Ck (j), 1, 1.5, 2, etc. are [wj] desirable as a value of R and S. This norm R and S does not need to be the same value. dik means distance of the k-th code vector to the i-th sound-source subvector. the code determination means 1007 -- etc. (several 11) etc. -- in computed distance, a code vector used as the minimum is elected and the index is coded as the code sequence 304. For example, when diu of inside with two or more above dik is the minimum, an index to the i-th subvector

coded is set to u. In the remainder creating means 1008, (several 12) generates the remainder subvector 1010 using a code vector elected by the code determination means 1007.

[0049]

[Mathematical formula 12]

$$res_i(j) = subvector_i(j) - C_u(j)$$

Here, the j-th element of the i-th remainder subvector 1010 is resi (j), and sets to Cu (j) the j-th element of the code vector elected by the code determination means 1007. The reversal process using the above-mentioned remainder subvector 1010 (several 10) is calculated, a vector is searched for, and the difference of this vector and the vector which was the coding subjects of the coding equipment concerned from the first is held as a MDCT coefficient which is the quantization target of each coding equipment after it. However, when coding of a certain zone is coding to the zone which does not affect the coding equipment after it (i.e., when subsequent coding equipment does not code), generation of remainder subvector [by the remainder creating means 1008] 1010 and MDCT1011 is unnecessary. Although any number of number of the code vector which the code book 1009 has is good, when memory space, computation time, etc. are taken into consideration, it is preferred to use about 64.

[0050]As other examples of the above-mentioned vector quantizer 1005, the following composition is also possible. That is, in the distance calculation means 1006, distance is computed using (several 13).

[0051]

[Mathematical formula 13]

$$dik = \begin{cases} \sum_{j=0}^{CR-1} w_j{}^R (subvector_i(j) - C_k(j))^S & k < K \\ \sum_{j=0}^{CR-1} w_j{}^R (subvector_i(j) - C_{K-k}(j))^S & k \geq K \end{cases}$$

However, K is a total of the code vector used for code search of the code book 1009. In the code determination means 1007, k which gives the minimum of the distance dik computed by (several 13) is elected, and the index is coded. However, k becomes a value to 0 to 2K-1. In the remainder creating means 1008, the remainder subvector 1010 is generated using (several 14).

[0052]

[Mathematical formula 14]

$$res_i(j) = \begin{cases} subvector_i(j) - C_u(j) & 0 \leq k < K \\ subvector_i(j) + C_u(j) & K \leq k < 2K \end{cases}$$

Although any number of number of the code vector which the code book 1009 has is good here, when the capacity of a memory, computation time, etc. are taken into consideration, it is preferred to use about 64. Although the composition which generates the dignity subvector 1004 only from the normalization ingredient 1002 was described above, it is also possible to multiply the dignity subvector 1004 by the dignity in consideration of human being's aural characteristic further, and to generate a dignity subvector. The bandwidth of each coding equipment of two or more stages of each, the number of coding equipment, and a connection order foreword are determined dynamically as mentioned above. And it quantizes based on the information on each coding equipment determined in this way.

[0053]A normalization code sequence which is an output of coding equipment of each zone in the decoding device 1002 on the other hand, It decodes using a code sequence from a quantizing part corresponding to this normalization code sequence, a bandwidth control code sequence which is the outputs of a coding band

control part in coding equipment further, and an analysis length code sequence which is the outputs of an analysis length judgment part.
[0054]Composition of the decoding machines 1202 and 1203 and -- is shown in drawing 9. Each decoding machine consists of the reverse normalizing part 1102 which carries out the multiplication of the inverse quantization part 1101 which reproduces a normalized MDCT coefficient, a MDCT coefficient which decoded a normalization coefficient and was reproduced [above-mentioned], and which were normalized, and the normalization coefficient.
[0055]In the reverse normalizing part 1102, from the normalization code sequence 303 from the normalizing part 301 of each coding equipment, a parameter used for normalization with the coding equipment 1 is restored, the multiplication of an output and this parameter of the inverse quantization part 1101 is carried out, and a MDCT coefficient is restored.
[0056]In the decoding band control part 1201, the bandwidth control code sequence 508 which is an output of the coding band control part 507 is used, Information on arrangement of coding equipment used with coding equipment and the number of coding equipment is restored, Based on the information, each decoding machines 1202, 1203, 1204, and 1202b are arranged to each zone, and a MDCT coefficient is obtained by the band composition part 9 which compounds a zone to a reverse order with encoding order of each coding equipment 511, 512, 513, and 511b in coding equipment. In this way, in the frequency time converter 5 which considers this obtained MDCT coefficient as an input, reverse MDCT is performed and restoration to a signal of a segment of time of a frequency domain from a signal is performed. Calculation of the above-mentioned reverse MDCT coefficient is shown by (several 15), for example.
[0057]
[Mathematical formula 15]

$$xx(n) = \frac{2}{N} \sum_{k=0}^{N-1} yy_k \cos\left\{\frac{2\pi(k+1/2)(n+n_0)}{N}\right\}$$

$$n_0 = \frac{N}{4} + \frac{1}{2}$$

Here, yyk is the MDCT coefficient restored in the band composition part 9, and xx (n) is a reverse MDCT coefficient and considers this as the output of the frequency time converter 5. In the window credit part 6, window credit is performed using output xx(i) from the frequency time converter 5. Window credit performs processing shown by (several 16), using the window used in the window credit part 202 in the temporal modulation converter 503 of the coding equipment 1.
[0058]
[Mathematical formula 16]

$$z(i) = xx(i)*h_i$$

Here, z(i) is an output of the window credit part 6. In the frame superposition part 7, an audio signal is reproduced using the output from the window credit part 6. Since the output from the window credit part 6 serves as a signal duplicate in time, it is made into the output signal of the decoding device 1002 in the frame superposition part 7, for example using (several 17).
[0059]
[Mathematical formula 17]

$$out_m(i) = z_m(i) + z_{m-1}(I+SHIFT)$$

zm(i) is output signal z(i) of the i-th window credit part 6 of the m-th time frame, and here zm-1(i), the -- considering it as the output signal of the i-th window credit part 6 of m-1 time frame, the measurement size and out m(i) equivalent to the analysis length 504 of coding equipment make SHIFT the output signal of the

decoding device 1002 in the m-th time frame of the frame superposition part 7.
[0060]In this Embodiment 1, the analysis length 504 may restrict as follows the frequency range which is computed by the bandwidth calculation part 901 and which can be quantized in the coding band control part 507. For example, when the analysis length 504 is 256, the minimum of the frequency range of each coding equipment which can be quantized shall be about 4 kHz, and a maximum shall be about 24 kHz. When analysis length is 1024 or 2048, a minimum shall be 0 Hz and a maximum shall be about 16 kHz. Once the analysis length 504 is furthermore set to 256, it is also controllable by the quantization sequence decision section 902 to fix after that the frequency range of each quantizer which can be quantized, and arrangement of a quantizer between fixed time (for example, for about 20 msec). By using this processing, arrangement of a quantizer is fixed temporally and it is a feeling of receipts and payments of an audibility zone (like [when a zone with that suddenly low whose zone high till a certain moment was a main sound changes to a main sound]). It can control that feeling which had receipts and payments of a voice band occurs.
[0061]In an audio signal encoding device by such this Embodiment 1, and a decoding device. A characteristic judgment part which determines a frequency band of an audio signal which coding equipment of two or more stages of each quantizes, A frequency band determined by the above-mentioned characteristic judgment part and an audio signal from the first by which frequency conversion was carried out are considered as the input, Since it had composition which determines connection order of coding equipment of two or more above-mentioned stages of each, is provided with a quantization zone of coding equipment, and a coding band control part which changes connection order into a code sequence, and performs scalable coding accommodative, Also when coding various audio signals, an audio signal encoding device which performs quality and efficient adaptation scalable coding which can demonstrate sufficient performance, and a decoding device which decodes this can be obtained.
[0062](Embodiment 2) Drawing 14 thru/or drawing 20 are used and explained to drawing 14 about the embodiment of the invention 2. Drawing 14 shows a block diagram of the coding equipment 2001 by the embodiment of the invention 2 which performs adaptation scalable coding, and the decoding device 2002. As shown in a figure, in the coding equipment 2001 200105, The number of coding equipment, the bit rate, a sampling frequency of an input audio signal, Coding conditions, such as coding band information of each coding equipment, a characteristic judgment part which determines a frequency band of an audio signal which each coding equipment of each stage of plurality [200107] quantizes, A frequency band where 200109 was determined by coding zone arrangement information, and 200110 was determined by the characteristic judgment part 200107, Considering as an input an audio input signal by which frequency conversion was carried out, a coding row and 200112 are modulation-code-ized sequence composing devices a quantization zone of coding equipment of two or more above-mentioned stages of each and a coding band control part which changes connection order into a code sequence, and 200111.
[0063]In the decoding device 2002, a decoding band control part which controls a decoding zone of each decoding machine which 200151 considers it as a coding row by 200150 considering it as a modulation-code-ized sequence resolver, and 200153b considers the coding row 200151 as an input, and decrypts this, and 200154b are decoding spectra.
[0064]As well as the above-mentioned Embodiment 1, although the coding equipment 2001 by the embodiment of the invention 2 performs adaptation scalable coding, Compare with Embodiment 1 and the coding band control part 200110 which contains the decoding band control part 200153 in the coding equipment 2001 newly, The decoding band control part 200153b which performs the same processing as the above-mentioned decoding band control part 200153 is added to the decoding device 2002, In the characteristic judgment part 200107 of this Embodiment 2, As it replaces with the spectrum power calculation part 803 of the characteristic judgment part 506 in the above-mentioned Embodiment 1 and is shown in drawing 16, Form the acoustic-sense mental model calculation part 200602, and in this characteristic judgment part 200107 further The coding condition 200105, The coding zone arrangement

information creating means 200604 which generates the coding zone arrangement information 200109 is established from the coding band information 200702 calculated from the coding zone calculation part 200601, and the zone number 200606 outputted from the arrangement deciding part 200603.

[0065]In the decoding device 2002, a decoding band control part which controls a decoding zone of each decoding machine which 200151 considers it as a coding row by 200150 considering it as a modulation-code-ized sequence resolver, and 200153b considers the coding row 200151 as an input, and decrypts this, and 200154b are decoding spectra.

[0066]Next, operation of this Embodiment 2 is explained. In this Embodiment 2, the original audio signal 501 which it is going to code presupposes that it is a digital signal series which continues in time like the above-mentioned Embodiment 1. First, the spectrum 505 of an original audio signal is acquired by the same processing as the above-mentioned Embodiment 1. According to this Embodiment 2, the coding condition 200105 including the number of coding equipment, the bit rate, a sampling frequency of an input audio signal, and coding band information of each encoder is inputted into the characteristic judgment part 200107 in this coding equipment 2001 to the coding equipment 2001. The characteristic judgment part 200107 outputs a quantization zone of each coding equipment of two or more stages of each, the number, and the coding zone arrangement information 200109 including information on connection order, and inputs this into the coding band control part 200110. As shown in drawing 17, in the coding band control part 200110 in addition to coding zone arrangement information 200109, The spectrum 505 of an original audio signal is inputted and the coding row 200111 which coded with each coding equipment controlled by this coding band control part 200110 based on these is outputted, This is inputted into the modulation-code-ized sequence composing device 200112, and is compounded by this, and the compounded output is further transmitted to the decoding device 2002.

[0067]In the decoding device 2002, an output of the modulation-code-ized sequence composing device 200112 of the coding equipment 2001 is received by the modulation-code-ized sequence resolver 200150, and it decomposes into the coding row 200151 and the analysis length code sequence 200152. The coding row 200151 is inputted into the decoding band control part 200153b, and acquires the decoding spectrum 200154b decrypted with each decoding machine controlled by this decoding band control part. And the decoded signal 8 is acquired from this decoding spectrum 200154b and the analysis length coding row 200152 which is the outputs of the above-mentioned modulation-code-ized sequence resolver 200150 like the above-mentioned Embodiment 1 using the frequency time converter 5, the window credit part 6, and the frame superposition part 7.

[0068]Next, operation of the characteristic judgment part 200107 is explained using drawing 15 - drawing 20. The coding condition 200105 is used for this characteristic judgment part 200107. Spectral information, such as the coding zone calculation part 200601 which computes the coding zone arrangement information 200702, the spectrum 505 of an original audio signal, and the difference spectrum 200108, And from the coding band information 200702, the acoustic-sense mental model calculation part 200602 and the analysis length 503 which compute the auditory weights 200605 based on human being's acoustic-sense mental model are referred to, The arrangement deciding part 200603 which performs a weighting to the auditory weights 200605 further according to this, opts for arrangement of a zone of each coding equipment, and outputs the zone number 200606, and the coding condition 200105, It comprises the coding zone arrangement information creating means 200604 which generates the coding zone arrangement information 200109 from the coding band information 200702 calculated from the coding zone calculation part 200601, and the zone number 200606 outputted from the arrangement deciding part 200603.

[0069]The coding condition 200105 set up before the coding equipment 2001 starts operation is used for the coding zone calculation part 200601, Maximum fpu (k) of a coding zone which the coding equipment 2003 shown in drawing 15 codes The minimum fpl (k) is computed and it is sent to the coding zone arrangement information creating means 200604 as the coding band information 200702. Here, k is a number for treating

a coding zone, and it shows a zone with big frequency as k is set to pmax which is the maximum number set up beforehand from 0. An example of pmax is 4. An example of operation of the coding zone calculation part 200601 is shown in Table 2.
[0070]
[Table 2]
符号化条件：サンプリング周波数が48kHz,合計ビットレートが24kbpsの時

帯域 k	fpu (k)	fpl (k)
0	221	0
1	318	222
2	415	319
3	512	416

符号化条件：サンプリング周波数が24kHz,合計ビットレートが24kbpsの時

帯域 k	fpu (k)	fpl (k)
0	443	0
1	637	444
2	831	638
3	1024	832

The acoustic-sense mental model calculation part 200602 An output signal from the filter 701, Or based on human being's acoustic-sense mental model, the auditory weights 200605 are computed from spectral information, such as the difference spectrum 200108 which is an output of the coding band control part 200110, and the coding band information 200702 which is the outputs of the coding zone calculation part 200601. An important zone is a big value on an acoustic sense, and these auditory weights 200605 seem to become a value with a small zone which is not so important on an acoustic sense. As an example of the acoustic-sense mental model calculation part 200602, there is a thing using a method of calculating power of an input spectrum. They are auditory weights when a spectrum inputted is made into x602(i). wpsy(k),
[0071]
[Mathematical formula 18]

$$w_{psy}(k) = \sum_{i=f_{pl}(k)}^{f_{pu}(k)} \left\{ x_{602}(i)^2 * \frac{1}{f_{pu}(k) - f_{pl}(k)} \right\}$$

It becomes. In this way, the computed auditory weights 200605, It is inputted into the arrangement deciding part 200603, and in this arrangement deciding part 200603. When the analysis length 503 is smallness, for example, 128, referring to the analysis length 503, So that the auditory weights 200605 of the zone of 4, whose zone number 200606 is size may become large, For example, this zone number doubles the weighting of the auditory weights of the zone of 4, and when the analysis length 503 is not smallness. The auditory weights 200605 are left intact, and these auditory weights 200603 calculate the zone used as the maximum, and send the zone number 200606 to the coding zone arrangement information creating means 200604.
[0072]The coding zone arrangement information creating means 200604 are the above-mentioned coding band information 200702 and the zone number 200606, and a thing that outputs the coding zone arrangement information 200109 for the coding condition 200105 as an input further. Namely, this coding zone arrangement information creating means 200604, While the coding zone arrangement information 200109 is needed considering this coding condition always referring to the coding condition 200105, The coding zone arrangement information 200109 which connects the above-mentioned coding band information 200702 and the zone number 200606 is outputted, and if this is required and is lost, operation which stops

the output will be carried out. For example, the zone number 200606 is outputted until it becomes the number of coding equipment specified by the coding condition 200105. in addition -- in the above-mentioned arrangement deciding part 200603 -- the analysis length 503 -- smallness -- sometimes, the zone number 200606 to output may be fixed.

[0073]Next, operation of the coding band control part 200110 is explained using drawing 17. The coding band control part 200110 considers the coding zone arrangement information 200109 which is an output from the above-mentioned characteristic judgment part 200107, and the spectrum 505 of an original audio signal as an input, Consider the coding row 200111 and the difference spectrum 200108 as the output, and to the inside. Receive the coding zone arrangement information 200109 and The spectrum 505 of an original audio signal, and the spectrum 505 of this past original audio signal, The difference spectrum 200108 with the spectrum 200705 which coded and decrypted this spectrum 505, the difference calculating means 200703 which takes difference with the spectral shift means 200701, the coding equipment 2003, and the spectrum 505 of the above-mentioned field audio signal and the decoding spectrum 200705 which are shifted to the zone of the zone number 200606, and the difference spectrum holding mechanism 200704 -- and, The synthetic spectrum 2001001 which decoded the code sequence 200111 with the decoding machine 2004, Based on the coding zone arrangement information 200702, a spectral shift is performed, this is compounded one by one, a synthetic spectrum is acquired, and the decoding band control part 200153 which computes the decoding spectrum 2007056 is included. Although the composition of the spectral shift means 200701 is as being shown in drawing 20, it uses the former spectrum 2001101 to shift and the coding zone arrangement information 200109 as an input. The spectrum 2001101 to shift among inputs of the spectral shift means 200701 in the coding band control part 200110, It is the spectrum 505 or the difference spectrum 200108 of an original audio signal, and they are shifted to the zone of the zone number 200606, and the shifted spectrum 2001102 and the coding band information 200702 of the coding zone arrangement information 200109 are outputted. The zone corresponding to the zone number 200606 can be searched for from fpl (k) of the coding band information 200702, and fpu (k). The procedure to shift is moving the spectrum between the above fpl (k) and fpu (k) to the zone which the coding equipment 2003 can process.

[0074]In this way, the coding equipment 2003 which considers the shifted spectrum 2001102 as an input, As shown in drawing 15, output the normalization code sequence 303 and the remainder code sequence 304, and Them, What united the coding band information 200702 which is an output of the spectral shift means 200701 is sent to the modulation-code-ized composing device 200112 and the decoding band control part 200153 as the code sequence 200111.

[0075]The above-mentioned coding row 200111 which is an output of the above-mentioned coding equipment 2003 is inputted into the decoding band control part 200153 in this coding band control part 200110. The operation of this decoding band control part 200153 is the same as what exists in the decoding device 2002 (200153b).

[0076]Next, the composition of the decoding band control part 200153b which exists in the above-mentioned decoding device 2002 is shown in drawing 19. It is what the decoding band control part 200153b considers the code sequence 200111 from the modulation-code-ized sequence resolver 200150 as an input, and outputs the decoding spectrum 200705b, It has the decoding machine 2004, the spectral shift means 200701, and the decoding spectrum calculation part 2001003 in the inside.

[0077]Composition of the above-mentioned decoding machine 2004 is shown in drawing 18. The decoding machine 2004 comprises the inverse quantization part 1101 and the reverse normalizing part 1102, and the inverse quantization part 1101, This remainder code sequence 304 is changed into a code index by considering the remainder code sequence 304 as an input among the code sequences 200111, and the code is reproduced with reference to a code book used with the coding equipment 2003. A reproduced code is sent to the reverse normalizing part 1102, and multiplication is carried out to the normalization system sequence of numbers 303a reproduced from the normalization code sequence 303 within the code sequence 200111,

and it acquires the synthetic spectrum 2001001. This synthetic spectrum 2001001 is inputted into the spectral shift means 200701.
[0078]Although an output of the decoding band control part 200153 in the coding band control part 200110 serves as the decoding spectrum 200705, this is the same as the decoding spectrum 200705b which is an output of the decoding band control part 200153b in the decoding device 2002.
[0079]The synthetic spectrum 2001001 compounded with the decoding machine 2004 is shifted by the spectral shift means 200701, the shifted synthetic spectrum 2001002 is acquired, and this is inputted into the decoding spectrum calculation part 2001003.
[0080]Within the decoding spectrum calculation part 2001003, the spectrum which holds the inputted synthetic spectrum and is held, and the newest synthetic spectrum are added, and operation outputted as the decoding spectrum 200705b is carried out.
[0081]The difference calculating means 200703 in the coding band control part 200110 calculates the difference of the spectrum 505 of an original audio signal, and the decoding spectrum 200705, the difference spectrum 200108 is outputted, and this is fed back to the characteristic judgment part 200107. Simultaneously, the above-mentioned difference spectrum 200108 is held by the difference spectrum holding mechanism 200704, is sent also to the spectral shift means 200701, and it is constituted so that it may have, when the following coding zone arrangement information 200109 is inputted. At the characteristic judgment part 200107, referring to a coding condition, outputting the coding zone arrangement information 200109 is continued until it fulfills this coding condition, and operation of the coding band control part 200110 is also suspended in the stage whose it was lost. The above-mentioned coding band control part 200110 has the difference spectrum holding mechanism 200704, in order to calculate the difference spectrum 200108. This is a storage area required in order to hold a difference spectrum, and is the arrangement which can memorize 2048 numbers, for example.
[0082]So that the coding condition 200105 may be fulfilled As mentioned above, the characteristic judgment part 200107, One by one, the coding row 200111 is outputted, it is sent to the modulation-code-ized sequence composing device 200112, and with the analysis length code sequence 510, processing by the coding band control part 200110 following it is repeated, and it is transmitted [it is compounded as a modulation-code-ized sequence and] to the decoding device 2002.
[0083]In the decoding device 2002, the modulation-code-ized sequence resolver 200150 decomposes into the coding row 200151 and the analysis length code sequence 200152 the modulation-code-ized sequence transmitted from the coding equipment 2001. This coding row 200151 and the analysis length code sequence 200152 are the same as the coding row 200111 in the coding equipment 2001, and the analysis length code sequence 510.
[0084]In the decoding band control part 200153b, the decomposed coding row 200151 is changed into the decoding spectrum 200154b, and this decoding spectrum 200154b, Using the information on the analysis length code sequence 200152, it is changed into the signal of a segment of time in the frequency time converter 5, the window credit part 6, and the frame superposition part 7, and it serves as the decoded signal 8.
[0085]Thus, according to the audio signal encoding device by this Embodiment 2, and the decoding device. The characteristic judgment part which determines the frequency band of the audio signal which the coding equipment of two or more stages of each quantizes like the above-mentioned Embodiment 1, The frequency band determined by the above-mentioned characteristic judgment part and the audio signal from the first by which frequency conversion was carried out are considered as the input, In the composition which determines the connection order of the coding equipment of two or more above-mentioned stages of each, is provided with the quantization zone of coding equipment, and the coding band control part which changes connection order into a code sequence, and performs scalable coding accommodative, While providing a decoding band control part in a decoding device, the coding band control part which contains a decoding

band control part in coding equipment, Since the spectrum power calculation part in a characteristic judgment part was made into the acoustic-sense mental model calculation part and it had further composition which established the coding zone arrangement information creating means in this characteristic judgment part, By having changed and replaced with the spectrum power calculation part of a characteristic judgment part, and having used the acoustic-sense mental model calculation part, an important portion can be auditorily judged with sufficient accuracy, and the zone can be chosen more. If a coding condition is fulfilled by the target audio signal encoding device [this invention] and a decoding device when performing the operation which opts for arrangement of coding equipment, processing of coding will be judged to be O.K. and coding zone arrangement information will not come out, either, but. In the operation for determining, arrangement of this coding equipment in the above-mentioned Embodiment 1. Each bandwidth when choosing a zone when arranging coding equipment, and the dignity of each zone to being immobilization in this Embodiment 2. As criteria of a characteristic judgment part, the sampling frequency of an input signal, and a compression ratio, Namely, since it is as the bit rate of coding, and ** ON, according to these, can change the weighting degree to each zone when choosing zone arrangement of each above-mentioned coding equipment, and further as criteria of a characteristic judgment part, When the conditions of the compression ratio are also contained, when a compression ratio is high (i.e., when the bit rate is low), make it not make the weighting degree of each zone when choosing zone arrangement of each above-mentioned coding equipment change it not much, and When a compression ratio is low on the other hand, When the bit rate is high, in order to pursue efficiency more, the weighting degree of each zone when choosing zone arrangement of each above-mentioned coding equipment Namely, on an acoustic sense, A more important place is emphasized and, thereby, the best balance of a compression ratio and quality can be obtained. Thus, also when coding various audio signals, the audio signal encoding and the decoding device which demonstrate sufficient performance and perform quality and efficient adaptation scalable coding can be obtained.

[0086]

[Effect of the Invention]As mentioned above, according to the audio signal encoding method concerning this invention, and the audio signal decoding method, a coding step, Have two or more coding sub steps, control of a coding bandwidth control step performs multistage coding of an audio signal, output encoded information, and a characteristic determination step, Output the zone weight information which judges the audio signal inputted and shows the weighting of each frequency band to code, and a coding bandwidth control step, The quantization zone of each coding sub step which constitutes multistage coding based on zone weight information, The multistage coding which determines connection order and is constituted scalable based on the quantization zone of each determined coding sub step and connection order is made to perform to a coding step, To the audio signal which has various character by having made it output the bandwidth control code sequence which shows the quantization zone of each determined coding sub step, and connection order, more by high-quality sound. An advantageous effect [say / that more efficient adaptation scalable coding can be performed] is acquired.

[Translation done.]

* NOTICES *

TECHNICAL FIELD

[Field of the Invention]This invention about an audio signal encoding method and an audio signal decoding method, The changed signal is compared with an original audio signal using the characteristic quantity especially obtained from audio signals, such as an audio signal and a music signal, especially the signal which changed the audio signal into the frequency domain from the segment of time using techniques, such as an orthogonal transformation, High quality and a broadband audio signal are related with the decoding method of the composition which can be decoded only using the methods of coding efficiently, in order to express by the fewest possible code sequences, all the coding rows that are the coded signals, or its part.

[Translation done.]

PRIOR ART

[Description of the Prior Art]The various techniques of coding and decrypting an audio signal efficiently are proposed. In the compression encoding type of the audio signal which has not less than 20-kHz frequency bands, such as a music signal, it is [an MPEG audio system and] Twin VQ. There are a system (TC-WVQ) etc. The coding mode represented by the MPEG system is a system which changes the digital audio signals of a time-axis into the data on a frequency axis using orthogonal transformations, such as cosine transformation, and codes the information on the frequency axis from important information auditorily using human being's auditory sensitivity characteristic.
Unimportant information and redundant information are systems which are not coded auditorily.
On the other hand, it is Twin VQ. A system (TC-WVQ) has a coding mode which it is going to express for the quite small amount of information to the amount of information of an original digital signal using the vector quantization technique. An MPEG audio and Twin VQ (TC-WVQ), Respectively ISO/IEC standard IS-11172-3 and T.Moriya, H.Suga:An 8 Kbits transform coder for noisy channels, and Proc.ICASSP. It is stated to 89, pp196-199, etc.
[0003]Here, the general outline of a Twin VQ system is explained using drawing 10. The original audio signal 101 is inputted into the analysis length judgment part 102, and analysis length is computed. Simultaneously, the analysis length judgment part 102 quantizes the analysis length 112, and outputs the analysis length code sequence 111. Next, according to the analysis length 112, the original audio signal 101 is changed into the original audio signal 104 of a frequency domain by the temporal modulation converter 103. Next, the normalizing process (flattening processing) of the original audio signal 104 of a frequency domain is carried out by the normalization processing part (flattening treating part) 106, and it obtains the audio signal 108 after a normalizing process. A normalizing process calculates the frequency facies 105 from the original audio signal 104, and is performed by breaking by the frequency facies 105 which computed the original audio signal 104. The normalization processing part 106 quantizes the frequency facies information that it used for the normalizing process, and outputs the normalization code sequence 107. Next, the audio signal 108 after a normalizing process is quantized by the vector quantization part 109, and the code sequence 110 is acquired.
[0004]Even if it uses a part of code sequence inputted into a decoder in recent years, there is a thing with the structure which can reproduce an audio signal. The above-mentioned structure is called scalable structure and it calls it scalable coding to code so that scalable structure can be realized.
[0005]An example of fixed scalable coding adopted as drawing 11 by the general Twin VQ system is shown. According to the analysis length 1314 judged by the analysis length judgment part 1303, the original audio signal 1304 of a frequency domain is obtained from the original audio signal 1301 by the temporal modulation converter 1302. Next, if the original audio signal 1304 of a frequency domain is inputted into the low-pass coding equipment 1305, the quantization error 1306 and the low-pass code sequence 1311 will be outputted. If the quantization error 1306 is inputted into the mid-range coding equipment 1307, the

quantization error 1308 and the mid-range code sequence 1312 will be outputted. If the quantization error 1308 is inputted into the high region coding equipment 1309, the quantization error 1310 and the high region code sequence 1313 will be outputted. Here the above-mentioned low-pass one, a mid-range, or high region coding equipment, Having a normalization processing part and a vector quantization part, the output outputs low-pass and a mid-range including each code sequence outputted by the quantization error, the normalization processing part, and the vector quantization part, or a high region code sequence.

[Translation done.]

EFFECT OF THE INVENTION

[Effect of the Invention]As mentioned above, according to the audio signal encoding method concerning this invention, and the audio signal decoding method, a coding step, Have two or more coding sub steps, control of a coding bandwidth control step performs multistage coding of an audio signal, output encoded information, and a characteristic determination step, Output the zone weight information which judges the audio signal inputted and shows the weighting of each frequency band to code, and a coding bandwidth control step, The quantization zone of each coding sub step which constitutes multistage coding based on zone weight information, The multistage coding which determines connection order and is constituted scalable based on the quantization zone of each determined coding sub step and connection order is made to perform to a coding step, To the audio signal which has various character by having made it output the bandwidth control code sequence which shows the quantization zone of each determined coding sub step, and connection order, more by high-quality sound. An advantageous effect [say / that more efficient adaptation scalable coding can be performed] is acquired.

[Translation done.]

TECHNICAL PROBLEM

[Problem to be solved by the invention]Since low-pass, a mid-range, and each zone quantizer of the high region were being fixed in fixed scalable coding of the conventional system as shown in drawing 11, as shown in drawing 12, it was difficult to code to distribution of an original audio signal, so that a quantization error may be lessened as much as possible. So, when coding an audio signal with various character and distribution, it was difficult to be unable to demonstrate sufficient performance but to perform efficient scalable coding by high-quality sound.
[0007]When it was made in order that this invention might cancel the above-mentioned problem, and coding an audio signal when coding various audio signals, as shown in drawing 13, It aims at providing the audio signal encoding method which it is efficient, and is a low bit rate, and can code to high-quality sound, and an audio signal decoding method by carrying out scalable coding of the various audio signals accommodative.

[Translation done.]

* NOTICES *

MEANS

[Means for solving problem]In order to solve this SUBJECT, an audio signal encoding method concerning this invention, and an audio signal decoding method, Not using fixed scalable coding, it is made to perform adaptation scalable coding to which character of an original audio signal and a frequency range coded in accordance with distribution are changed.
[0009]An audio signal encoding method concerning this invention A characteristic determination step, a coding bandwidth control step, A coding step is included and it is time. - Are an audio signal by which frequency conversion was carried out an audio signal encoding method changed into a coding row, and a coding row, Including encoded information and a bandwidth control code sequence, a coding step, Have two or more coding sub steps, control of a coding bandwidth control step performs multistage coding of an audio signal, output encoded information, and a characteristic determination step, Output zone weight information which judges an audio signal inputted and shows a weighting of each frequency band to code, and a coding bandwidth control step, A quantization zone of each coding sub step which constitutes multistage coding based on zone weight information, Connection order is determined, multistage coding constituted scalable based on a quantization zone of each determined coding sub step and connection order is made to perform to a coding step, and a bandwidth control code sequence which shows a quantization zone of each determined coding sub step and connection order is outputted.
[0010]In said audio signal encoding method, the audio signal encoding method concerning this invention determines the quantization zone of each coding sub step, and connection order so that a coding bandwidth control step may become either of the multistage coding defined beforehand.
[0011]In said audio signal encoding method the audio signal encoding method concerning this invention, A coding step outputs a quantization error and a coding bandwidth control step determines the quantization zone of each coding sub step, and connection order based on zone weight information and a quantization error.
[0012]The audio signal decoding method concerning this invention A decoding bandwidth control step, Are the coding row which includes a decoding step and includes encoded information and a bandwidth control code sequence an audio signal decoding method decoded to an audio signal, and a bandwidth control code sequence, The quantization zone of each coding at the time of carrying out the multistage coding of the encoded information and connection order are shown, and a decoding step, It has two or more decoding sub steps, control of a decoding bandwidth control step performs a multistage decoding of encoded information, and it is made for a decoding bandwidth control step to make the multistage decoding constituted scalable based on a bandwidth control code sequence perform to a decoding step.
[0013]
[Mode for carrying out the invention]Hereafter, the embodiment of the invention 1 is described about Embodiment 2 using drawing 14 thru/or drawing 20, using drawing 1 thru/or drawing 9.
[0014](Embodiment 1) Drawing 1 shows the block diagram of an audio signal encoding device by the

embodiment of the invention 1 which performs adaptation scalable coding. In drawing 1, 1001 is coding equipment which codes the original audio signal 501. The analysis length judgment part which judges the analysis length 504 at the time of 502 analyzing the above-mentioned field audio signal 501 in this coding equipment 1001, and 503 are the units of the above-mentioned analysis length 504, The temporal modulation converter which changes the time-axis of the original audio signal 501 into a frequency axis, The analysis length by which 504 was judged by the above-mentioned analysis length judgment part 502, and 505 The spectrum of an original audio signal, The filter into which the spectrum 505 of this original audio signal is inputted 701, Each coding equipment 511,512,513,511b of two or more stages of each which 506 judges the characteristic of the spectrum 505 of an original audio signal, and can be set to the above-mentioned coding equipment 1001, The characteristic judgment part which determines the frequency band of the audio signal which ********, and the frequency band of each coding equipment in which 507 was determined by this characteristic judgment part 506, Consider the audio signal by which frequency conversion was carried out [above-mentioned] as the input, and Each coding equipment 512,513,514,511b of two or more stages of each, The bandwidth control code sequence which is the above-mentioned code sequence which determines ********, and to which the quantization zone of each coding equipment and the coding band control part which changes connection order into a code sequence, and 508 are outputted from this coding band control part 507, The analysis length code sequence made into the code sequence and 511,512,513 the above-mentioned analysis length 504 to which 510 was outputted from the above-mentioned analysis length judgment part 502, The low-pass coding equipment which was mentioned above and which codes low-pass, a mid-range, and the signal of a high region, respectively, Mid-range coding equipment, high region coding equipment, the 2nd step low-pass coding equipment in which 511b codes the quantization error 518 of the low-pass coding equipment 511 of the 1st step, The low-pass code sequence which is a coded signal with which 521,522,523 is outputted from this each coding equipment 511,512,513, A mid-range code sequence, a high region code sequence, the 2nd step low-pass code sequence whose 521b is an encoding output of the 2nd step low-pass coding equipment 511b, The quantization error which is a difference of the signal before being coded and each above-mentioned coded signal with which 518,519,520 is outputted from this each coding equipment 511,512,513, and 518b are the 2nd step quantization errors which are quantization errors of the 2nd step low-pass coding equipment 511b.

[0015]On the other hand, 1002 is a decoding device which decrypts the coding row coded with the above-mentioned coding equipment 1001. The frequency time converter which performs conversion with 5 [contrary to the temporal modulation converter 503 in the above-mentioned coding equipment 1001] in this decoding device 1002, The window credit part which performs window credit which 6 multiplies by a windowing function on a time-axis, and 7 A frame superposition part, 8 a decoded signal and 9 a band composition part and 1201 a decoding band control part, and 1202-1203 and 1204, The low-pass decoding machine which decrypts corresponding to the above-mentioned low-pass coding equipment, mid-range coding equipment, and the high region coding equipment 511,512,513, respectively, a mid-range decoding machine, a high region decoding machine, and 1202b are the 2nd step low-pass decoding machines which decrypt the output of the 1st step low-pass decoding machine 1202.

[0016]here -- the zone of further others [machine / the coding equipment after the 2nd step, and / decoding] -- the accuracy of coding and a decoding can be improved if needed, so that it may provide also in multistage and this becomes multistage.

[0017]Hereafter, operation of the coding equipment 1001 is explained first. The original audio signal 501 which it is going to code presupposes that it is a digital signal series which continues in time. For example, suppose that it is an audio signal the digital signal quantized to 16 bits by 48 kHz of sampling frequencies.

[0018]The above-mentioned field audio signal 501 is inputted into the analysis length judgment part 502. The above-mentioned analysis length judgment part 502 judges the characteristic of the inputted above-mentioned field audio signal 501, it opts for the analysis length 504, and the result is sent to the decoding

device 1002 as the analysis length code sequence 510. As the analysis length 504, 256, 1024, 4096, etc. are used, for example. For example, when the high region frequency component contained in the original audio signal 501 exceeds a predetermined value. The analysis length 504 is set to 256, a low-pass frequency component exceeds a predetermined value, and when a high region frequency component is smaller than a predetermined value, the analysis length 504 is set to 4096, and when other, the analysis length 504 is set to 1024. In this way, according to the determined analysis length 504, the spectrum 505 of the original audio signal 501 is computed by the temporal modulation converter 503.

[0019]The block diagram of the temporal modulation converter 503 in the audio signal encoding device by the embodiment of the invention 1 is shown in drawing 2. The above-mentioned field audio signal 501 will output, if it is accumulated by the frame dividing part 201 and the this accumulated measurement size reaches the analysis length 504 determined by the above-mentioned analysis length judgment part 502 until the sampled value reaches a predetermined measurement size. In the case where the frame dividing part 201 is a thing of composition of outputting for every shift length of a certain, for example, the analysis length 504 is made into 4096 samples, If the shift length of the half of the analysis length 504 is set up, the analysis length 504 has the composition of outputting the 4096 newest samples for every time to be equivalent to reaching 2048 samples. Though natural, even if the analysis length 504 and a sampling frequency change, it is possible to have similarly the composition which set shift length as the half of the analysis length 504. And the output from this frame dividing part 201 is inputted into the latter window credit part 202. In the window credit part 202, to the output from the frame dividing part 201, it multiplies by a windowing function on a time-axis, and is considered as the output of the window credit part 202. This situation is shown by (several 1), for example.

[0020]

[Mathematical formula 1]

$$hxi = h_i * x_i \quad i = 1,2,\cdots,N$$

$$h_i = \sin\left\{\frac{\pi}{N}(i+0.5)\right\}$$

However, xi is an output from the frame dividing part 201, hi is a windowing function and hxi is an output from the window credit part 202 here. i is still a suffix of time. The windowing function hi shown by (several 1) needs to be an example, and the windowing function does not necessarily need to be a thing of (several 1).

[0021]It depends for selection of a windowing function on the feature of the signal inputted into the window credit part 202, the analysis length 504 of the frame dividing part 201, and the form of the windowing function in the frame located forward and backward in time. For example, as a feature of the signal inputted into the window credit part 202, when the analysis length 504 of the frame dividing part 201 is set to N, the average power of the signal inputted for every N/4 is computed, When changing the average power very sharply, it chooses performing the operation shown for making the analysis length 504 shorter than N (several 1) etc. Choosing suitably is desirable so that there may be no distortion in the form of the windowing function of the frame of the present time according to the form of the windowing function of the frame of the last time, and the form of the windowing function of a back frame.

[0022]Subsequently, the output from the window credit part 202 is inputted into the MDCT section 203, a modification discrete cosine transform is given here, and a MDCT coefficient is outputted. The general formula of a modification discrete cosine transform is expressed with (several 2).

[0023]

[Mathematical formula 2]

$$y_k = \sum_{n=0}^{N-1} hx_n * \cos\left\{\frac{2\pi\left(k+\frac{1}{2}\right)(n+n_0)}{N}\right\}$$

$$n_0 = \frac{N}{4} + \frac{1}{2} \quad \left(k = 0,1,\cdots,\frac{N}{2}-1\right)$$

Thus, the MDCT coefficient which is an output of the MDCT section 203, (Several 2) Supposing it can express with inner yk, the output of the MDCT section 203 shows a frequency characteristic, and it corresponds to a high frequency component at linearity, so that it increases from 0 to a low frequency component, so that its variable k of yk is [zero] near, and it becomes close to N/2-1. In this way, the computed above-mentioned MDCT coefficient serves as the spectrum 505 of an original audio signal.

[0024]Next, the spectrum 505 of the above-mentioned field audio signal is inputted into the filter 701. If this filter 701 is inputted as x701(i) and an output is made into y701(i), the filter expressed with (several 3) will be used, for example.

[0025]

[Mathematical formula 3]

$$y_{701}(i) = w_{701}(i) * \{x_{701}(i) + x_{701}(i+1)\}$$
$$i = 0,1,\cdots,fs-2$$

Here, fs is the analysis length 504. Although the filter 701 expressed with (several 3) is a kind of moving average filter, though natural, it is not necessary to limit it to a moving average filter, they may be other high pass filters, and may be a zone control filter.

[0026]The output of the filter 701 and the analysis length 504 which computed by the analysis length judgment part 502 are inputted into the characteristic judgment part 506. The details of the characteristic judgment part 506 are shown in drawing 6. In the characteristic judgment part 506, the auditory and physical characteristic of spectrum 505** of the original audio signal 501 and an original audio signal is determined. Auditory and the physical characteristic of the original audio signal 501 and this spectrum 505 are the differences between a sound, music, or **, for example. In the case of a sound, most frequency components are, for example on low-pass from 6 kHz.

[0027]Next, operation of the characteristic judgment part 506 is explained using drawing 6. If a signal which filtered the spectrum 505 of an original audio signal inputted into the characteristic judgment part 506 with the filter 701 is made into x506(i), Based on this x506(i), spectrum power p506(i) is calculated by the spectrum power calculation part 803 by (several 4).

[0028]

[Mathematical formula 4]

$$p_{506}(i) = x_{506}(i)^2$$

Set this spectrum power p506(i) to one of the inputs of the coding band control part 507, and let it be the bandwidth control dignity 517 of each coding equipment. When the analysis length 504 is small, for example it becomes 256, it determines by the arrangement deciding part 804, and the coding zone arrangement information 516 is sent to the coding band control part 507 as fixed location so that each coding equipment may be arranged fixed.

[0029]When the analysis length 504 is small and it is except, it determines by the arrangement deciding part 804 at the time of 4096 or 1024, and it sends the coding zone arrangement information 516 to the coding band control part 507 as dynamic allocation, for example so that each coding equipment may be arranged dynamically.

[0030]Next, operation of the coding band control part 507 is explained using drawing 7. The bandwidth

control dignity 517 which is an output from the above-mentioned characteristic judgment part 506 at the coding band control part 507, The signal which filtered the coding zone arrangement information 516 and the spectrum 505 of the original audio signal with the filter 701, the quantization error 518 which each coding equipment outputted, or 519 or 520 is inputted. However, there are these inputs, in order that each coding equipment 511, 512, 513, and 511b and the coding band control part 507 may operate recursively, and in operation of the first-time coding band control part 507, since there is no quantization error, it becomes three inputs except a quantization error.

[0031]As mentioned above, when the analysis length 504 is small and the coding zone arrangement information 516 is placed in a fixed position, According to fixed location of the zone defined beforehand, it codes by determining the quantization zone of coding equipment, the number, and connection order by the quantization sequence decision section 902 and the number deciding part 903 of coding equipment, and the bandwidth calculation part 901 so that coding may be performed from low-pass in order to a mid-range and a high region. That is, the band information, the number of coding equipment, and its connection order foreword of coding equipment are coded as information by the bandwidth control code sequence 508 at that time.

[0032]for example, the coding zone and the number of coding equipment of each coding equipment -- respectively -- it codes by arranging coding equipment to one and 0 Hz - 8 kHz, at 0 Hz - 4 kHz so that it may be set to one and 8 kHz - 16 kHz with two and may become three at one and 4 kHz - 12 kHz at 16 kHz - 24 kHz.

[0033]Next, operation of the coding band control part 507 when the coding zone arrangement information 516 is dynamic allocation is explained. The coding band control part 507 consists of the bandwidth calculation part 901 which determines the quantization bandwidth of each coding equipment, the quantization sequence decision section 902 which determines a quantization order of each coding equipment, and three of number deciding part of coding equipment 903** which determines the number of the coding equipment of each zone further. Although it is a translation which determines the bandwidth of each coding equipment based on the signal inputted into the coding band control part 507, In each zone (a predetermined zone, for example, 0 Hz - 4 kHz, 0 kHz - 8 kHz, 4 kHz - 12 kHz, 8 kHz - 16 kHz, and 16 kHz - 24 kHz), the bandwidth control dignity 517 and the average value of what carried out the multiplication of the quantization error after each coding equipment codes are computed. Here, if bandwidth control dignity 517 is made into weight517(i) and a quantization error is made into err507(i), average value will be computed by (several 5).

[0034]

[Mathematical formula 5]

$$Ave_{901}(j) = \frac{1}{f_{upper}(j) - f_{lower}(j)} \sum_{i=f_{upper}(j)}^{f_{lower}(j)} weight_{517}(i) * err_{507}(i)^2$$

Average value [in / j / here / in the index of each zone, and Ave901 (j) / the zone j], and fupper (j) And flower(s) (j) are the upper limited frequency of the zone j, and a lower cut off frequency. In this way, average value Ave901 (j) obtained search j used as the maximum, and it serves as a zone which coding equipment codes. Send the value of searched j to the number deciding part 903 of coding equipment, and the one number of coding equipment of the zone corresponding to j is increased, It memorizes how many coding equipment exist in a predetermined coding zone, and coding is repeated until the sum total of the memorized number of coding equipment becomes a total of the coding equipment determined beforehand. Finally, the zone and the number of coding equipment of coding equipment are transmitted to a decoding machine as the bandwidth control code sequence 508.

[0035]Next, operation of the coding equipment 3 is explained using drawing 3. The coding equipment 3

consists of the normalizing part 301 and the quantizing part 302. In the normalizing part 301, a MDCT coefficient is normalized using some parameters by considering both with the MDCT coefficient which are a signal of the time-axis which is an output from the frame dividing part 201, and an output from the MDCT section 203 as an input. Normalization of a MDCT coefficient means here oppressing dispersion in the size of the MDCT coefficient which has a difference in a size dramatically by the low-pass ingredient and a high-frequency component, For example, to a high-frequency component, a low-pass ingredient points out oppressing dispersion in the size of a MDCT coefficient of a low-pass ingredient by electing a parameter which serves as a big value and a value small in a high-frequency component, and doing division of the above-mentioned MDCT coefficient now, when very large. In the normalizing part 301, the index expressing the parameter used for normalization is coded as the normalization code sequence 303.

[0036]In the quantizing part 302, the MDCT quantization of coefficient is performed by considering the MDCT coefficient normalized by the normalizing part 301 as an input. Under the present circumstances, this quantizing part 302 outputs this such code IDDEKKUSU to which the difference between the this quantized value and each quantization output corresponding to two or more code indices in a code book becomes the smallest. In this case, the difference of the value quantized by the above-mentioned quantizing part 302 and the value corresponding to the code index outputted from this quantizing part 302 is a quantization error.

[0037]Next, a detailed example of the above-mentioned normalizing part 301 is explained using drawing 4. In drawing 4, the frequency facies normalizing part in which 401 receives the output of the frame dividing part 201 and the MDCT section 203, and 402 are zone amplitude normalizing parts which undergo the output of the above-mentioned frequency facies normalizing part 401, and normalize with reference to the band table 403.

[0038]Next, operation is explained. In the frequency facies normalizing part 401, using data output on a time-axis from the frame dividing part 201, frequency facies which are facies of rough frequency are computed, and division of the MDCT coefficient which is an output from the MDCT section 203 is done. A parameter used for expressing frequency facies is coded as the normalization code sequence 303. In the zone amplitude normalizing part 402, it normalizes for every zone shown with the band table 403 by considering an output signal from the frequency facies normalizing part 401 as an input. for example, -- supposing a MDCT coefficient which is an output of the frequency facies normalizing part 401 considers it as dct (i) (i = 0-2047) and the band table 403 is a thing as shown in (Table 1), for example -- etc. (several 6) etc. -- average value of amplitude for every zone is computed by using.

[0039]

[Table 1]

帯域 k	$f_{lower}(k)$	$f_{upper}(k)$
0	0	10
1	11	22
2	23	33
3	34	45
4	46	56
5	57	68
6	69	80
7	81	92
8	93	104
9	105	116
10	117	128
11	129	141
12	142	153
13	154	166
14	167	179
15	180	192
16	193	205
17	206	219
18	220	233
19	234	247
20	248	261
21	262	276
22	277	291
23	292	307
24	308	323
25	324	339
26	340	356
27	357	374
28	375	392
29	393	410
30	411	430
31	431	450
32	451	470
33	471	492
34	493	515
35	516	538
36	539	563
37	564	587
38	589	615
39	616	643
40	645	673
41	674	705
42	706	737
43	738	772
44	773	809
45	810	848
46	849	889
47	890	932
48	933	978
49	979	1027
50	1028	1079

帯域 k	$f_{lower}(k)$	$f_{upper}(k)$
51	1080	1135
52	1136	1193
53	1194	1255
54	1256	1320
55	1321	1389
56	1390	1462
57	1463	1538
58	1539	1617
59	1618	1699
60	1700	1783
61	1784	1870
62	1871	1958
63	1959	2048

[Mathematical formula 6]

$$sum_j = \sum_{i=bjlow}^{bjhigh} dct(i)^p$$

$$ave_j = \left(\frac{sum_j}{bjhigh - bjlow + 1} \right)^{-p} \quad bjlow \leq i \leq bjhigh$$

Here, bjlow and bjhigh show most the low-pass index i and the index i of most a high region whose dct(i) in the j-th zone shown in the band table 403 belongs, respectively. p is the norm in distance calculation and 2 etc. is desirable. avej is the average value of the amplitude in each zone number j. In the zone amplitude normalizing part 402, qavej is computed by quantizing avej, for example, it normalizes using (several 7).

[0040]
[Mathematical formula 7]

$$n\_dct(i) = dct(i) / gave_j \qquad bjlow \leq i \leq bjhigh$$

Quantization of avej may use quantization of a scalar and may perform vector quantization using a code book. In the zone amplitude normalizing part 402, the index of the parameter used for expressing qavej is coded as the normalization code sequence 303.
[0041]Although the composition of the normalizing part 301 in coding equipment showed the thing of composition of having used both the frequency facies normalizing part 401 of drawing 4 and the zone amplitude normalizing part 402, the composition only using the frequency facies normalizing part 401 may be used for it, and the composition only using the zone amplitude normalizing part 402 may be used for it. It is good also as composition which inputs the output signal of the MDCT section 203 into the quantizing part 302 as it is with the composition which does not use above-mentioned both by the low-pass ingredient of the MDCT coefficient outputted from the MDCT section 203, and a high-frequency component when there is no big dispersion.
[0042]Next, the details of the frequency facies normalizing part 401 of drawing 4 are explained using drawing 5. In drawing 5, the linear-predictive-coding part in which 601 receives the output of the frame dividing part 201, the facies quantizing part in which 602 receives the output of the linear-predictive-coding part 601, and 603 are envelopment characteristic normalizing parts which undergo the output of the MDCT section 203.
[0043]Next, operation of the above-mentioned frequency facies normalizing part 401 is explained with reference to drawing 5. In the above-mentioned linear-predictive-coding part 601, linear predictive coding (Linear Predictive Coding) is conducted by considering the audio signal on the time-axis from the frame dividing part 201 as an input. The linear predictor coefficients (LPC coefficient) of linear predictive coding compute the autocorrelation function of signals by which window credit was carried out, such as a humming window, are solving a normal equation etc. and, generally can compute it. The computed linear predictor coefficients are changed into a line spectrum pair coefficient (LSP (LineSpectrum Pair) coefficient) etc., and are quantized by the facies quantizing part 602. As the quantization technique here, vector quantization may be used and scalar quantity child-ization may be used. And the frequency transfer characteristic which the parameter quantized by the facies quantizing part 602 expresses is computed by the envelopment characteristic normalizing part 603, and the MDCT coefficient which is an output from the MDCT section 203 is normalized by doing division now. If linear predictor coefficients equivalent to the parameter quantized by the facies quantizing part 602 as a concrete example of calculation are made into qlpc(i), the above-mentioned frequency transfer characteristic computed by the envelopment characteristic normalizing part 603 can be expressed with (several 8), for example.
[0044]
[Mathematical formula 8]

$$li = \begin{cases} qlpc(i) & 0 \leq i \leq ORDER \\ 0 & ORDER+1 \leq i \leq N \end{cases}$$

$$env(i) = \frac{1}{fft(li)}$$

Here, as for ORDER, ten to about 40 are desirable. fft() means Fast Fourier Transform. using computed frequency transfer characteristic env(i) -- the envelopment characteristic normalizing part 603 -- the following -- being shown (several 9) -- it normalizes by using.
[Mathematical formula 9]

$$\mathrm{fdct}(i) = \frac{\mathrm{mdct}(i)}{\mathrm{env}(i)}$$

Here, mdct(i) is an output signal from the MDCT section 203, and fdct(i) is an output signal from the normalized envelopment characteristic normalizing part 603.
[0045]Next, detailed operation of the quantization method of the quantizing part 302 in the above-mentioned coding equipment 1 is explained using drawing 8. The MDCT coefficient 1001 inputted into the quantizing part 302 extracts some from the MDCT coefficient 1001, and constitutes the sound-source subvector 1003. Similarly when the coefficient sequence which divided the MDCT coefficient which is an input of the normalizing part 301 by the MDCT coefficient which is an output of the normalizing part 301 is used as the normalization ingredient 1002, by the normalizing part 301 also about this normalization ingredient 1002. A subvector can be extracted from this normalization ingredient 1002, and the dignity subvector 1004 can consist of same rules as having extracted the sound-source subvector 1003 from the MDCT coefficient 1001. The rule which extracts the sound-source subvector 1003 and the dignity subvector 1004 from the MDCT coefficient 1001 and the normalization ingredient 1002, respectively has a method etc. which are shown by (several 10), for example.
[0046]
[Mathematical formula 10]

$$\mathrm{subvector}_i(j) = \begin{cases} \mathrm{vector}\left(\frac{\mathrm{VTOTAL}}{\mathrm{CR}} \cdot i + j\right) & \text{ただし} \frac{\mathrm{VTOTAL}}{\mathrm{CR}} * i + j < \mathrm{TOTAL} \\ 0 & \text{ただし} \frac{\mathrm{VTOTAL}}{\mathrm{CR}} * i + j \geq \mathrm{TOTAL} \end{cases}$$

The j-th element of the i-th sound-source subvector is subvector i (j), the MDCT coefficient 1001 is vector() here, and the total number of elements of the MDCT coefficient 1001 is TOTAL, The number of elements of the sound-source subvector 1003 is the value as TOTAL with same CR and VTOTAL, or a larger value, and it sets up so that VTOTAL/CR may become a positive number value. For example, when TOTAL is 2048, CR is 19, VTOTAL is 2052, CR is 23, VTOTAL is 2070, CR is 21, and VTOTAL is 2079 etc. The dignity subvector 1004 can also be extracted in several 10 procedure. In the vector quantizer 1005, out of a code vector in the code book 1009. An index of a code vector with which distance with the sound-source subvector 1003 looked for that to which weighting **** also becomes small by the dignity subvector 1004, and gave the minimum distance, The remainder subvector 1010 equivalent to a quantization error of a code vector and the input sound source subvector 1003 which gave the minimum distance is outputted.
[0047]In a actual example of a computational procedure, the vector quantizer 1005 explains as what consists of three components, the distance calculation means 1006, the code determination means 1007, and remainder creating means 1008**. In the distance calculation means 1006, distance of the i-th sound-source subvector 1003 and the k-th code vector of the code book 1009 is computed, for example using (several 11).
[0048]
[Mathematical formula 11]

$$\mathrm{dik} = \sum_{j=0}^{CR-1} w_j{}^{R} (\mathrm{subvector}_i(j) - C_k(j))^S$$

Here, the j-th element of the k-th code vector, R, and S are the norm of distance calculation, and, as for the j-th element of a dignity subvector, and Ck (j), 1, 1.5, 2, etc. are [wj] desirable as a value of R and S. This norm R and S does not need to be the same value. dik means distance of the k-th code vector to the i-th sound-source subvector. the code determination means 1007 -- etc. (several 11) etc. -- in computed distance, a code vector used as the minimum is elected and the index is coded as the code sequence 304. For

example, when diu of inside with two or more above dik is the minimum, an index to the i-th subvector coded is set to u. In the remainder creating means 1008, (several 12) generates the remainder subvector 1010 using a code vector elected by the code determination means 1007.
[0049]
[Mathematical formula 12]
$$res_i(j) = subvector_i(j) - C_u(j)$$

Here, the j-th element of the i-th remainder subvector 1010 is resi (j), and sets to Cu (j) the j-th element of a code vector elected by the code determination means 1007. A reversal process using the above-mentioned remainder subvector 1010 (several 10) is calculated, a vector is searched for, and a difference of this vector and a vector which was the coding subjects of the coding equipment concerned from the first is held as a MDCT coefficient which is the quantization target of each coding equipment after it. However, when coding of a certain zone is coding to a zone which does not affect coding equipment after it (i.e., when subsequent coding equipment does not code), generation of remainder subvector [by the remainder creating means 1008] 1010 and MDCT1011 is unnecessary. Although any number of number of a code vector which the code book 1009 has is good, when memory space, computation time, etc. are taken into consideration, it is preferred to use about 64.
[0050]As other examples of the above-mentioned vector quantizer 1005, the following composition is also possible. That is, in the distance calculation means 1006, distance is computed using (several 13).
[0051]
[Mathematical formula 13]
$$dik = \begin{cases} \sum_{j=0}^{CR-1} w_j{}^R (subvector_i(j) - C_k(j))^S & k < K \\ \sum_{j=0}^{CR-1} w_j{}^R (subvector_i(j) - C_{K-k}(j))^S & k \geq K \end{cases}$$

However, K is a total of a code vector used for code search of the code book 1009. In the code determination means 1007, k which gives the minimum of the distance dik computed by (several 13) is elected, and the index is coded. However, k becomes a value to 0 to 2K-1. In the remainder creating means 1008, the remainder subvector 1010 is generated using (several 14).
[0052]
[Mathematical formula 14]
$$res_i(j) = \begin{cases} subvector_i(j) - C_u(j) & 0 \leq k < K \\ subvector_i(j) + C_u(j) & K \leq k < 2K \end{cases}$$

Although any number of number of a code vector which the code book 1009 has is good here, when capacity of a memory, computation time, etc. are taken into consideration, it is preferred to use about 64. Although composition which generates the dignity subvector 1004 only from the normalization ingredient 1002 was described above, it is also possible to multiply the dignity subvector 1004 by dignity in consideration of human being's aural characteristic further, and to generate a dignity subvector. Bandwidth of each coding equipment of two or more stages of each, the number of coding equipment, and a connection order foreword are determined dynamically as mentioned above. And it quantizes based on information on each coding equipment determined in this way.
[0053]The normalization code sequence which is an output of the coding equipment of each zone in the decoding device 1002 on the other hand, It decodes using the code sequence from the quantizing part

corresponding to this normalization code sequence, the bandwidth control code sequence which is the outputs of the coding band control part in coding equipment further, and the analysis length code sequence which is the outputs of an analysis length judgment part.
[0054]The composition of the decoding machines 1202 and 1203 and -- is shown in drawing 9. Each decoding machine consists of the reverse normalizing part 1102 which carries out the multiplication of the inverse quantization part 1101 which reproduces the normalized MDCT coefficient, the MDCT coefficient which decoded the normalization coefficient and was reproduced [above-mentioned], and which were normalized, and the normalization coefficient.
[0055]In the reverse normalizing part 1102, from the normalization code sequence 303 from the normalizing part 301 of each coding equipment, the parameter used for normalization with the coding equipment 1 is restored, the multiplication of the output and this parameter of the inverse quantization part 1101 is carried out, and a MDCT coefficient is restored.
[0056]In the decoding band control part 1201, the bandwidth control code sequence 508 which is an output of the coding band control part 507 is used, The information on arrangement of the coding equipment used with coding equipment and the number of coding equipment is restored, Based on the information, each decoding machines 1202, 1203, 1204, and 1202b are arranged to each zone, and a MDCT coefficient is obtained by the band composition part 9 which compounds a zone to a reverse order with the encoding order of each coding equipment 511, 512, 513, and 511b in coding equipment. In this way, in the frequency time converter 5 which considers this obtained MDCT coefficient as an input, reverse MDCT is performed and restoration to the signal of a segment of time of a frequency domain from a signal is performed. Calculation of the above-mentioned reverse MDCT coefficient is shown by (several 15), for example.
[0057]
[Mathematical formula 15]

$$xx(n) = \frac{2}{N} \sum_{k=0}^{N-1} yy_k \cos\left\{ \frac{2\pi(k+1/2)(n+n_0)}{N} \right\}$$

$$n_0 = \frac{N}{4} + \frac{1}{2}$$

Here, yyk is the MDCT coefficient restored in the band composition part 9, and xx (n) is a reverse MDCT coefficient and considers this as the output of the frequency time converter 5. In the window credit part 6, window credit is performed using output xx(i) from the frequency time converter 5. Window credit performs processing shown by (several 16), using the window used in the window credit part 202 in the temporal modulation converter 503 of the coding equipment 1.
[0058]
[Mathematical formula 16]

$$z(i) = xx(i)*h_i$$

Here, z(i) is an output of the window credit part 6. In the frame superposition part 7, an audio signal is reproduced using the output from the window credit part 6. Since the output from the window credit part 6 serves as a signal duplicate in time, it is made into the output signal of the decoding device 1002 in the frame superposition part 7, for example using (several 17).
[0059]
[Mathematical formula 17]

$$out_m(i) = z_m(i) + z_{m-1}(i + SHIFT)$$

zm(i) is output signal z(i) of the i-th window credit part 6 of the m-th time frame, and here zm-1(i), the -- considering it as the output signal of the i-th window credit part 6 of m-1 time frame, the measurement size

and out m(i) equivalent to the analysis length 504 of coding equipment make SHIFT the output signal of the decoding device 1002 in the m-th time frame of the frame superposition part 7.
[0060]In this Embodiment 1, the analysis length 504 may restrict as follows the frequency range which is computed by the bandwidth calculation part 901 and which can be quantized in the coding band control part 507. For example, when the analysis length 504 is 256, the minimum of the frequency range of each coding equipment which can be quantized shall be about 4 kHz, and a maximum shall be about 24 kHz. When analysis length is 1024 or 2048, a minimum shall be 0 Hz and a maximum shall be about 16 kHz. Once the analysis length 504 is furthermore set to 256, it is also controllable by the quantization sequence decision section 902 to fix after that the frequency range of each quantizer which can be quantized, and arrangement of a quantizer between fixed time (for example, for about 20 msec). By using this processing, arrangement of a quantizer is fixed temporally and it is a feeling of receipts and payments of an audibility zone (like [when a zone with that suddenly low whose zone high till a certain moment was a main sound changes to a main sound]). It can control that feeling which had receipts and payments of a voice band occurs.
[0061]In the audio signal encoding device by such this Embodiment 1, and a decoding device. The characteristic judgment part which determines the frequency band of the audio signal which the coding equipment of two or more stages of each quantizes, The frequency band determined by the above-mentioned characteristic judgment part and the audio signal from the first by which frequency conversion was carried out are considered as the input, Since it had composition which determines the connection order of the coding equipment of two or more above-mentioned stages of each, is provided with the quantization zone of coding equipment, and the coding band control part which changes connection order into a code sequence, and performs scalable coding accommodative, Also when coding various audio signals, the audio signal encoding device which performs quality and efficient adaptation scalable coding which can demonstrate sufficient performance, and the decoding device which decodes this can be obtained.
[0062](Embodiment 2) Drawing 14 thru/or drawing 20 are used and explained to drawing 14 about the embodiment of the invention 2. Drawing 14 shows the block diagram of the coding equipment 2001 by the embodiment of the invention 2 which performs adaptation scalable coding, and the decoding device 2002. As shown in a figure, in the coding equipment 2001 200105, The number of coding equipment, the bit rate, the sampling frequency of an input audio signal, Coding conditions, such as coding band information of each coding equipment, the characteristic judgment part which determines the frequency band of the audio signal which each coding equipment of each stage of plurality [200107] quantizes, The frequency band where 200109 was determined by coding zone arrangement information, and 200110 was determined by the characteristic judgment part 200107, Considering as an input the audio input signal by which frequency conversion was carried out, a coding row and 200112 are modulation-code-ized sequence composing devices the quantization zone of the coding equipment of two or more above-mentioned stages of each and the coding band control part which changes connection order into a code sequence, and 200111.
[0063]In the decoding device 2002, the decoding band control part which controls the decoding zone of each decoding machine which 200151 considers it as a coding row by 200150 considering it as a modulation-code-ized sequence resolver, and 200153b considers the coding row 200151 as an input, and decrypts this, and 200154b are decoding spectra.
[0064]As well as the above-mentioned Embodiment 1, although the coding equipment 2001 by the embodiment of the invention 2 performs adaptation scalable coding, Compare with Embodiment 1 and the coding band control part 200110 which contains the decoding band control part 200153 in the coding equipment 2001 newly, The decoding band control part 200153b which performs the same processing as the above-mentioned decoding band control part 200153 is added to the decoding device 2002, In the characteristic judgment part 200107 of this Embodiment 2, As it replaces with the spectrum power calculation part 803 of the characteristic judgment part 506 in the above-mentioned Embodiment 1 and is shown in drawing 16, Form the acoustic-sense mental model calculation part 200602, and in this

characteristic judgment part 200107 further The coding condition 200105, The coding zone arrangement information creating means 200604 which generates the coding zone arrangement information 200109 is established from the coding band information 200702 calculated from the coding zone calculation part 200601, and the zone number 200606 outputted from the arrangement deciding part 200603.

[0065]In the decoding device 2002, the decoding band control part which controls the decoding zone of each decoding machine which 200151 considers it as a coding row by 200150 considering it as a modulation-code-ized sequence resolver, and 200153b considers the coding row 200151 as an input, and decrypts this, and 200154b are decoding spectra.

[0066]Next, operation of this Embodiment 2 is explained. In this Embodiment 2, the original audio signal 501 which it is going to code presupposes that it is a digital signal series which continues in time like the above-mentioned Embodiment 1. First, the spectrum 505 of an original audio signal is acquired by the same processing as the above-mentioned Embodiment 1. According to this Embodiment 2, the coding condition 200105 including the number of coding equipment, the bit rate, the sampling frequency of an input audio signal, and the coding band information of each encoder is inputted into the characteristic judgment part 200107 in this coding equipment 2001 to the coding equipment 2001. The characteristic judgment part 200107 outputs the quantization zone of each coding equipment of two or more stages of each, the number, and the coding zone arrangement information 200109 including the information on connection order, and inputs this into the coding band control part 200110. As shown in drawing 17, in the coding band control part 200110 in addition to coding zone arrangement information 200109, The spectrum 505 of an original audio signal is inputted and the coding row 200111 which coded with each coding equipment controlled by this coding band control part 200110 based on these is outputted, This is inputted into the modulation-code-ized sequence composing device 200112, and is compounded by this, and the compounded output is further transmitted to the decoding device 2002.

[0067]In the decoding device 2002, the output of the modulation-code-ized sequence composing device 200112 of the coding equipment 2001 is received by the modulation-code-ized sequence resolver 200150, and it decomposes into the coding row 200151 and the analysis length code sequence 200152. The coding row 200151 is inputted into the decoding band control part 200153b, and acquires the decoding spectrum 200154b decrypted with each decoding machine controlled by this decoding band control part. And the decoded signal 8 is acquired from this decoding spectrum 200154b and the analysis length coding row 200152 which is the outputs of the above-mentioned modulation-code-ized sequence resolver 200150 like the above-mentioned Embodiment 1 using the frequency time converter 5, the window credit part 6, and the frame superposition part 7.

[0068]Next, operation of the characteristic judgment part 200107 is explained using drawing 15 - drawing 20. The coding condition 200105 is used for this characteristic judgment part 200107. Spectral information, such as the coding zone calculation part 200601 which computes the coding zone arrangement information 200702, the spectrum 505 of an original audio signal, and the difference spectrum 200108, And from the coding band information 200702, the acoustic-sense mental model calculation part 200602 and the analysis length 503 which compute the auditory weights 200605 based on human being's acoustic-sense mental model are referred to, The arrangement deciding part 200603 which performs a weighting to the auditory weights 200605 further according to this, opts for arrangement of the zone of each coding equipment, and outputs the zone number 200606, and the coding condition 200105, It comprises the coding zone arrangement information creating means 200604 which generates the coding zone arrangement information 200109 from the coding band information 200702 calculated from the coding zone calculation part 200601, and the zone number 200606 outputted from the arrangement deciding part 200603.

[0069]The coding condition 200105 set up before the coding equipment 2001 starts operation is used for the coding zone calculation part 200601, Maximum fpu (k) of the coding zone which the coding equipment 2003 shown in drawing 15 codes The minimum fpl (k) is computed and it is sent to the coding zone

arrangement information creating means 200604 as the coding band information 200702. Here, k is a number for treating a coding zone, and it shows the zone with big frequency as k is set to pmax which is the maximum number set up beforehand from 0. An example of pmax is 4. An example of operation of the coding zone calculation part 200601 is shown in Table 2.
[0070]
[Table 2]
符号化条件：サンプリング周波数が48kHz,合計ビットレートが24kbpsの時

帯域 k	fpu (k)	fpl (k)
0	221	0
1	318	222
2	415	319
3	512	416

符号化条件：サンプリング周波数が24kHz,合計ビットレートが24kbpsの時

帯域 k	fpu (k)	fpl (k)
0	443	0
1	637	444
2	831	638
3	1024	832

The acoustic-sense mental model calculation part 200602 The output signal from the filter 701, Or based on human being's acoustic-sense mental model, the auditory weights 200605 are computed from spectral information, such as the difference spectrum 200108 which is an output of the coding band control part 200110, and the coding band information 200702 which is the outputs of the coding zone calculation part 200601. An important zone is a big value on an acoustic sense, and these auditory weights 200605 seem to become a value with a small zone which is not so important on an acoustic sense. As an example of the acoustic-sense mental model calculation part 200602, there is a thing using the method of calculating the power of an input spectrum. They are auditory weights when the spectrum inputted is made into x602(i). wpsy(k), [0071]
[Mathematical formula 18]

$$w_{psy}(k) = \sum_{i=f_{pl}(k)}^{f_{pu}(k)} \left\{ x_{602}(i)^2 * \frac{1}{f_{pu}(k) - f_{pl}(k)} \right\}$$

It becomes. In this way, the computed auditory weights 200605, It is inputted into the arrangement deciding part 200603, and in this arrangement deciding part 200603. When the analysis length 503 is smallness, for example, 128, referring to the analysis length 503, So that the auditory weights 200605 of the zone of 4, whose zone number 200606 is size may become large, For example, this zone number doubles the weighting of the auditory weights of the zone of 4, and when the analysis length 503 is not smallness. The auditory weights 200605 are left intact, and these auditory weights 200603 calculate the zone used as the maximum, and send the zone number 200606 to the coding zone arrangement information creating means 200604.
[0072]The coding zone arrangement information creating means 200604 are the above-mentioned coding band information 200702 and the zone number 200606, and a thing that outputs the coding zone arrangement information 200109 for the coding condition 200105 as an input further. Namely, this coding zone arrangement information creating means 200604, While the coding zone arrangement information 200109 is needed considering this coding condition always referring to the coding condition 200105, The coding zone arrangement information 200109 which connects the above-mentioned coding band information

200702 and the zone number 200606 is outputted, and if this is required and is lost, operation which stops the output will be carried out. For example, the zone number 200606 is outputted until it becomes the number of coding equipment specified by the coding condition 200105. in addition -- in the above-mentioned arrangement deciding part 200603 -- the analysis length 503 -- smallness -- sometimes, the zone number 200606 to output may be fixed.

[0073]Next, operation of the coding band control part 200110 is explained using drawing 17. The coding band control part 200110 considers the coding zone arrangement information 200109 which is an output from the above-mentioned characteristic judgment part 200107, and the spectrum 505 of an original audio signal as an input, Consider the coding row 200111 and the difference spectrum 200108 as the output, and to the inside. Receive the coding zone arrangement information 200109 and The spectrum 505 of an original audio signal, and the spectrum 505 of this past original audio signal, The difference spectrum 200108 with the spectrum 200705 which coded and decrypted this spectrum 505, the difference calculating means 200703 which takes difference with the spectral shift means 200701, the coding equipment 2003, and the spectrum 505 of the above-mentioned field audio signal and the decoding spectrum 200705 which are shifted to a zone of the zone number 200606, and the difference spectrum holding mechanism 200704 -- and, The synthetic spectrum 2001001 which decoded the code sequence 200111 with the decoding machine 2004, Based on the coding zone arrangement information 200702, a spectral shift is performed, this is compounded one by one, a synthetic spectrum is acquired, and the decoding band control part 200153 which computes the decoding spectrum 2007056 is included. Although composition of the spectral shift means 200701 is as being shown in drawing 20, it uses the former spectrum 2001101 to shift and the coding zone arrangement information 200109 as an input. The spectrum 2001101 to shift among inputs of the spectral shift means 200701 in the coding band control part 200110, It is the spectrum 505 or the difference spectrum 200108 of an original audio signal, and they are shifted to a zone of the zone number 200606, and the shifted spectrum 2001102 and the coding band information 200702 of the coding zone arrangement information 200109 are outputted. A zone corresponding to the zone number 200606 can be searched for from fpl (k) of the coding band information 200702, and fpu (k). A procedure to shift is moving a spectrum between the above fpl (k) and fpu (k) to a zone which the coding equipment 2003 can process.

[0074]In this way, the coding equipment 2003 which considers the shifted spectrum 2001102 as an input, As shown in drawing 15, output the normalization code sequence 303 and the remainder code sequence 304, and Them, What united the coding band information 200702 which is an output of the spectral shift means 200701 is sent to the modulation-code-ized composing device 200112 and the decoding band control part 200153 as the code sequence 200111.

[0075]The above-mentioned coding row 200111 which is an output of the above-mentioned coding equipment 2003 is inputted into the decoding band control part 200153 in this coding band control part 200110. The operation of this decoding band control part 200153 is the same as what exists in the decoding device 2002 (200153b).

[0076]Next, the composition of the decoding band control part 200153b which exists in the above-mentioned decoding device 2002 is shown in drawing 19. It is what the decoding band control part 200153b considers the code sequence 200111 from the modulation-code-ized sequence resolver 200150 as an input, and outputs the decoding spectrum 200705b, It has the decoding machine 2004, the spectral shift means 200701, and the decoding spectrum calculation part 2001003 in the inside.

[0077]The composition of the above-mentioned decoding machine 2004 is shown in drawing 18. The decoding machine 2004 comprises the inverse quantization part 1101 and the reverse normalizing part 1102, and the inverse quantization part 1101, This remainder code sequence 304 is changed into a code index by considering the remainder code sequence 304 as an input among the code sequences 200111, and the code is reproduced with reference to the code book used with the coding equipment 2003. The reproduced code is sent to the reverse normalizing part 1102, and multiplication is carried out to the normalization system

sequence of numbers 303a reproduced from the normalization code sequence 303 within the code sequence 200111, and it acquires the synthetic spectrum 2001001. This synthetic spectrum 2001001 is inputted into the spectral shift means 200701.

[0078]Although the output of the decoding band control part 200153 in the coding band control part 200110 serves as the decoding spectrum 200705, this is the same as the decoding spectrum 200705b which is an output of the decoding band control part 200153b in the decoding device 2002.

[0079]The synthetic spectrum 2001001 compounded with the decoding machine 2004 is shifted by the spectral shift means 200701, the shifted synthetic spectrum 2001002 is acquired, and this is inputted into the decoding spectrum calculation part 2001003.

[0080]Within the decoding spectrum calculation part 2001003, the spectrum which holds the inputted synthetic spectrum and is held, and the newest synthetic spectrum are added, and operation outputted as the decoding spectrum 200705b is carried out.

[0081]The difference calculating means 200703 in the coding band control part 200110 calculates the difference of the spectrum 505 of an original audio signal, and the decoding spectrum 200705, the difference spectrum 200108 is outputted, and this is fed back to the characteristic judgment part 200107. Simultaneously, the above-mentioned difference spectrum 200108 is held by the difference spectrum holding mechanism 200704, is sent also to the spectral shift means 200701, and it is constituted so that it may have, when the following coding zone arrangement information 200109 is inputted. At the characteristic judgment part 200107, referring to a coding condition, outputting the coding zone arrangement information 200109 is continued until it fulfills this coding condition, and operation of the coding band control part 200110 is also suspended in the stage whose it was lost. The above-mentioned coding band control part 200110 has the difference spectrum holding mechanism 200704, in order to calculate the difference spectrum 200108. This is a storage area required in order to hold a difference spectrum, and is the arrangement which can memorize 2048 numbers, for example.

[0082]So that the coding condition 200105 may be fulfilled As mentioned above, the characteristic judgment part 200107, One by one, the coding row 200111 is outputted, it is sent to the modulation-code-ized sequence composing device 200112, and with the analysis length code sequence 510, processing by the coding band control part 200110 following it is repeated, and it is transmitted [it is compounded as a modulation-code-ized sequence and] to the decoding device 2002.

[0083]In the decoding device 2002, the modulation-code-ized sequence resolver 200150 decomposes into the coding row 200151 and the analysis length code sequence 200152 a modulation-code-ized sequence transmitted from the coding equipment 2001. This coding row 200151 and the analysis length code sequence 200152 are the same as the coding row 200111 in the coding equipment 2001, and the analysis length code sequence 510.

[0084]In the decoding band control part 200153b, the decomposed coding row 200151 is changed into the decoding spectrum 200154b, and this decoding spectrum 200154b, Using information on the analysis length code sequence 200152, it is changed into a signal of a segment of time in the frequency time converter 5, the window credit part 6, and the frame superposition part 7, and it serves as the decoded signal 8.

[0085]Thus, according to an audio signal encoding device by this Embodiment 2, and the decoding device. A characteristic judgment part which determines a frequency band of an audio signal which coding equipment of two or more stages of each quantizes like the above-mentioned Embodiment 1, A frequency band determined by the above-mentioned characteristic judgment part and an audio signal from the first by which frequency conversion was carried out are considered as the input, In composition which determines connection order of coding equipment of two or more above-mentioned stages of each, is provided with a quantization zone of coding equipment, and a coding band control part which changes connection order into a code sequence, and performs scalable coding accommodative, While providing a decoding band control part in a decoding device, a coding band control part which contains a decoding band control part in coding

equipment, Since a spectrum power calculation part in a characteristic judgment part was made into an acoustic-sense mental model calculation part and it had further composition which established a coding zone arrangement information creating means in this characteristic judgment part, By having changed and replaced with a spectrum power calculation part of a characteristic judgment part, and having used an acoustic-sense mental model calculation part, an important portion can be auditorily judged with sufficient accuracy, and the zone can be chosen more. If a coding condition is fulfilled by the target audio signal encoding device [this invention] and a decoding device when performing an operation which opts for arrangement of coding equipment, processing of coding will be judged to be O.K. and coding zone arrangement information will not come out, either, but. In an operation for determining, arrangement of this coding equipment in the above-mentioned Embodiment 1. Each bandwidth when choosing a zone when arranging coding equipment, and dignity of each zone to being immobilization in this Embodiment 2. As criteria of a characteristic judgment part, a sampling frequency of an input signal, and a compression ratio, Namely, since it is as the bit rate of coding, and ** ON, according to these, can change a weighting degree to each zone when choosing zone arrangement of each above-mentioned coding equipment, and further as criteria of a characteristic judgment part, When conditions of a compression ratio are also contained, when a compression ratio is high (i.e., when the bit rate is low), make it not make a weighting degree of each zone when choosing zone arrangement of each above-mentioned coding equipment change it not much, and When a compression ratio is low on the other hand, When the bit rate is high, in order to pursue efficiency more, a weighting degree of each zone when choosing zone arrangement of each above-mentioned coding equipment Namely, on an acoustic sense, A more important place is emphasized and, thereby, best balance of a compression ratio and quality can be obtained. Thus, also when coding various audio signals, audio signal encoding and a decoding device which demonstrate sufficient performance and perform quality and efficient adaptation scalable coding can be obtained.

[Translation done.]

DESCRIPTION OF DRAWINGS

[Brief Description of the Drawings]
[Drawing 1]The block diagram of adaptation scalable coding in the audio signal encoding device by the embodiment of the invention 1
[Drawing 2]The figure showing the temporal modulation converter in the coding equipment of the above-mentioned Embodiment 1
[Drawing 3]The figure showing the coding equipment in the coding equipment of the above-mentioned Embodiment 1
[Drawing 4]The figure showing the normalizing part in the coding equipment of the above-mentioned Embodiment 1
[Drawing 5]The figure showing the frequency facies normalizing part in the coding equipment of the above-mentioned Embodiment 1
[Drawing 6]The figure showing the characteristic judgment part in the coding equipment of the above-mentioned Embodiment 1
[Drawing 7]The figure showing the coding band control part in the coding equipment of the above-mentioned Embodiment 1
[Drawing 8]The figure showing the quantizing part in the coding equipment of the above-mentioned Embodiment 1
[Drawing 9]The figure showing the decoding machine in the coding equipment of the above-mentioned Embodiment 1
[Drawing 10]The figure showing the outline of a general TwinVQ system
[Drawing 11]The figure showing a general TwinVQ scalable coding method
[Drawing 12]The figure showing the demerit of general fixed scalable coding
[Drawing 13]The figure showing the strong point of general adaptation scalable coding
[Drawing 14]The block diagram of adaptation scalable coding in the audio signal encoding device by the embodiment of the invention 2
[Drawing 15]The figure showing the coding equipment in the coding equipment of the above-mentioned Embodiment 2
[Drawing 16]The figure showing the characteristic judgment part in the coding equipment of the above-mentioned Embodiment 2
[Drawing 17]The figure showing the coding band control part in the coding equipment of the above-mentioned Embodiment 2
[Drawing 18]The figure showing the decoding machine in the coding equipment of the above-mentioned Embodiment 2
[Drawing 19]The figure showing the decoding band control part in the coding equipment of the above-mentioned Embodiment 2

[Drawing 20]The figure showing the spectral shift means in the coding equipment of the above-mentioned Embodiment 2
[Explanations of letters or numerals]
1 Coding equipment
2 Decoding device
501 Original audio signal
502 Analysis length judgment part
503 Temporal modulation converter
504 Analysis length
505 The spectrum of an original audio signal
506 Characteristic judgment part
507 Coding band control part
508 Bandwidth control code sequence
510 Analysis length code sequence
511 Low-pass coding equipment
512 Mid-range coding equipment
513 High region coding equipment
511b The 2nd step low-pass coding equipment
518,519,520,518b Quantization error
521 Low-pass code sequence
522 Mid-range code sequence
523 High region code sequence
521b The 2nd step low-pass code sequence
701 Filter
5 Frequency time converter
6 Window credit part
7 Frame superposition part
8 Decoded signal
9 Band composition part
1201 Decoding band control part
1202 Low-pass decoding machine
1203 Mid-range decoding machine
1204 High region decoding machine
1202b The 2nd step low-pass decoding machine
201 Frame dividing part
202 Window credit part
203 MDCT section
3 Coding equipment
301 Normalizing part
302 Quantizing part
303 Normalization code sequence
304 A code sequence
401 A frequency facies normalizing part
402 A zone amplitude normalizing part
403 A band table
601 A linear-predictive-coding part
602 A facies quantizing part

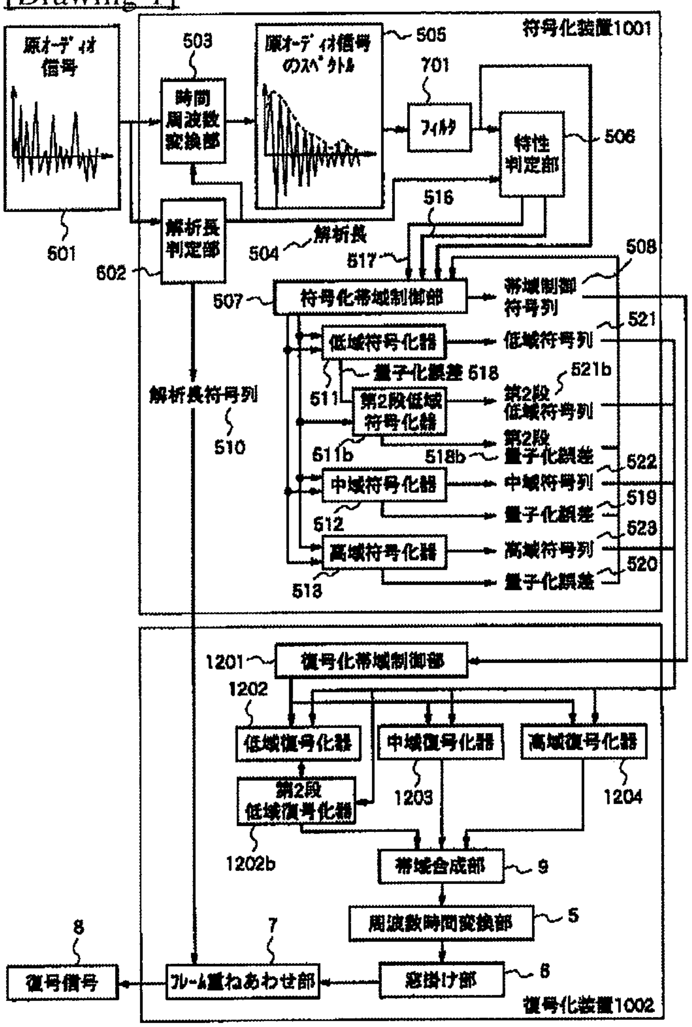
603 An envelopment characteristic normalizing part
803 A spectrum power calculation part
804 An arrangement deciding part
517 Bandwidth control dignity
516 Coding zone arrangement information
901 A bandwidth calculation part
902 A quantization sequence decision section
903 The number deciding part of coding equipment
1001 MDCT of a zone which a quantizing part quantizes
1002 A normalization ingredient of the same quantization zone
1003 A sound-source subvector
1004 A dignity subvector
1005 A vector quantizer
1006 A distance calculation means
1007 A code determination means
1008 A remainder creating means
1009 A code book
1010 A remainder subvector
1011 The remainder of MDCT of a zone which a certain quantizing part quantizes
101 An original audio signal
102 An analysis length judgment part
103 A temporal modulation converter
104 An original audio signal of a frequency domain
105 Frequency facies
106 A normalization processing part
107 A normalization code sequence
108 The present audio signal after a normalizing process
109 Vector quantization part
110 Code sequence
111 Analysis length code sequence
1301 Original audio signal
1302 Temporal modulation converter
1303 Analysis length judgment part
1304 The original audio signal of a frequency domain
1305 Low-pass coding equipment
1306 Quantization error
1307 Mid-range coding equipment
1308 Quantization error
1309 High region coding equipment
1310 Quantization error
1311 Low-pass code sequence
1312 Mid-range code sequence
1313 High region code sequence
1314 Analysis length code sequence
2001 Coding equipment
2002 Decoding device
200105 Coding condition

200107 Characteristic judgment part
200108 Difference spectrum
200109 Coding zone arrangement information
200110 Coding band control part
200111 Coding row
200112 Modulation-code-ized sequence composing device
200150 Modulation-code-ized sequence resolver
200151 Coding row
200152 Analysis length coding row
200153 Decoding band control part
200154 Decoding spectrum
2003 Coding equipment
200305 Coding band information
200601 A coding zone calculation part
200602 An acoustic-sense mental model calculation part
200603 An arrangement deciding part
200604 A coding zone arrangement information creating means
200605 Auditory weights
200701 A spectral shift means
200702 Coding band information
200703 A difference calculating means
200704 Difference spectrum holding mechanism
2004 A decoding machine
200901 A reverse quantification part
200902 A reverse normalizing part
2001001 A synthetic spectrum
2001002 A shifted synthetic spectrum
2001003 A decoding spectrum calculation part
2001101 A former spectrum
2001102 A shifted spectrum

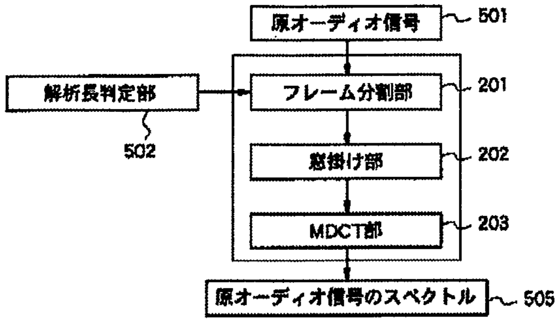
[Translation done.]

DRAWINGS

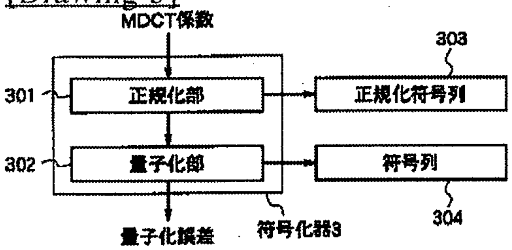
[Drawing 1]



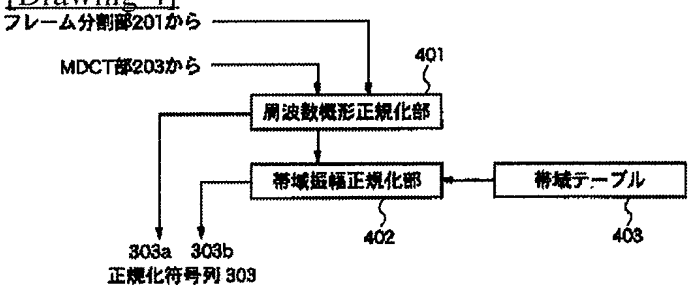
[Drawing 2]



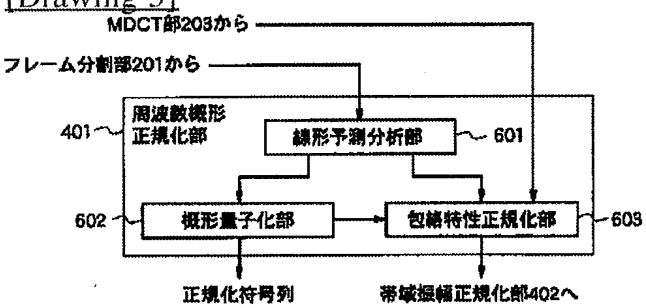
[Drawing 3]



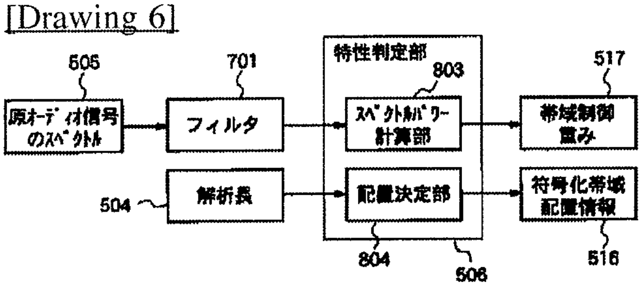
[Drawing 4]



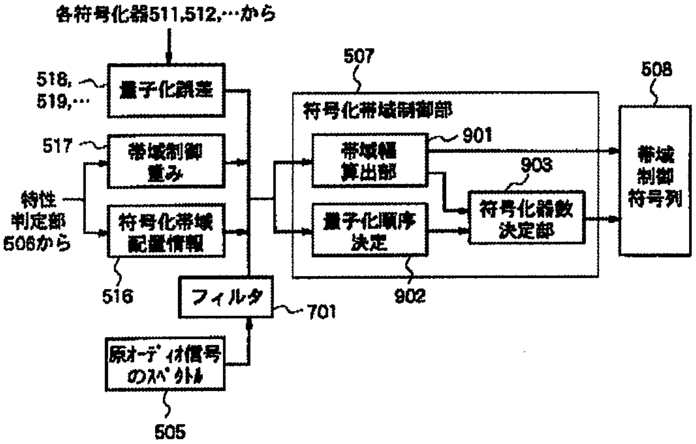
[Drawing 5]



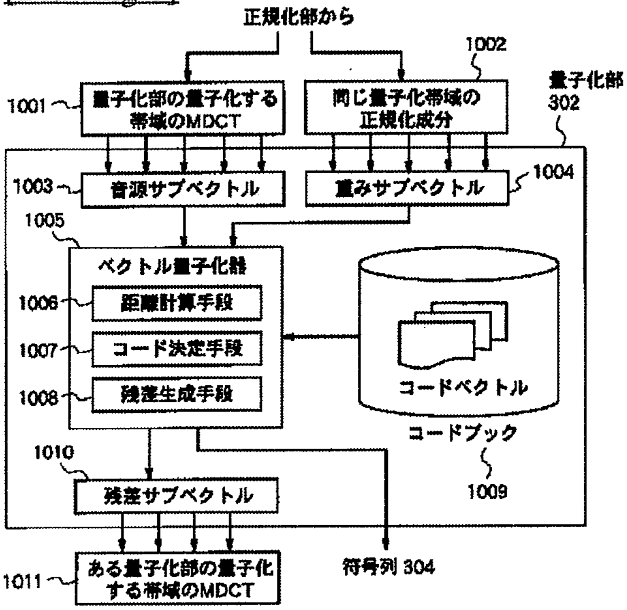
[Drawing 6]



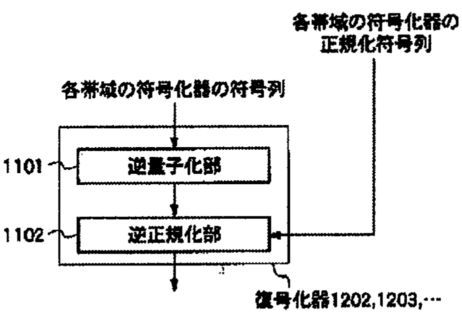
[Drawing 7]



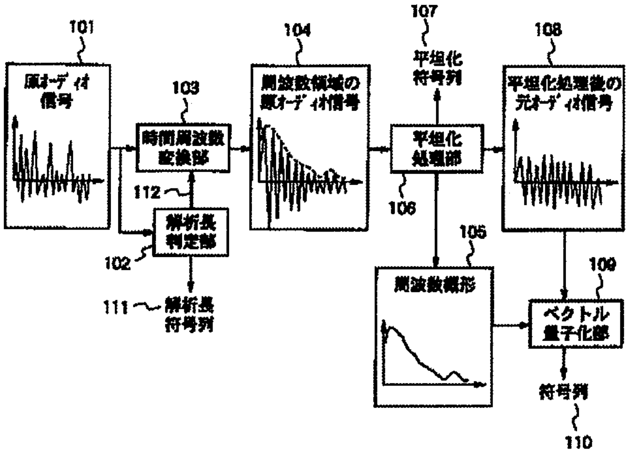
[Drawing 8]



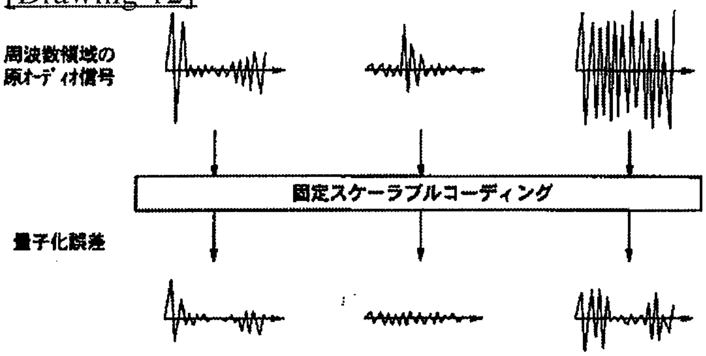
[Drawing 9]



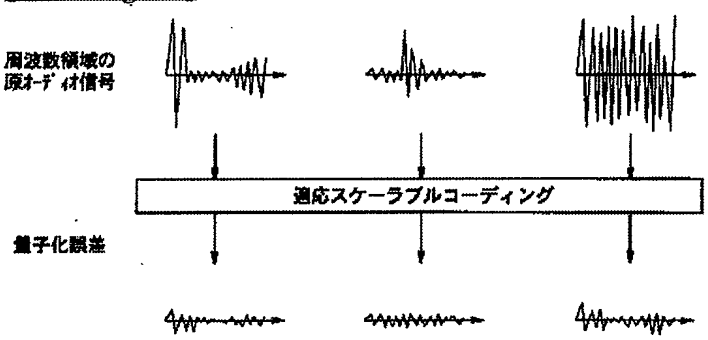
[Drawing 10]



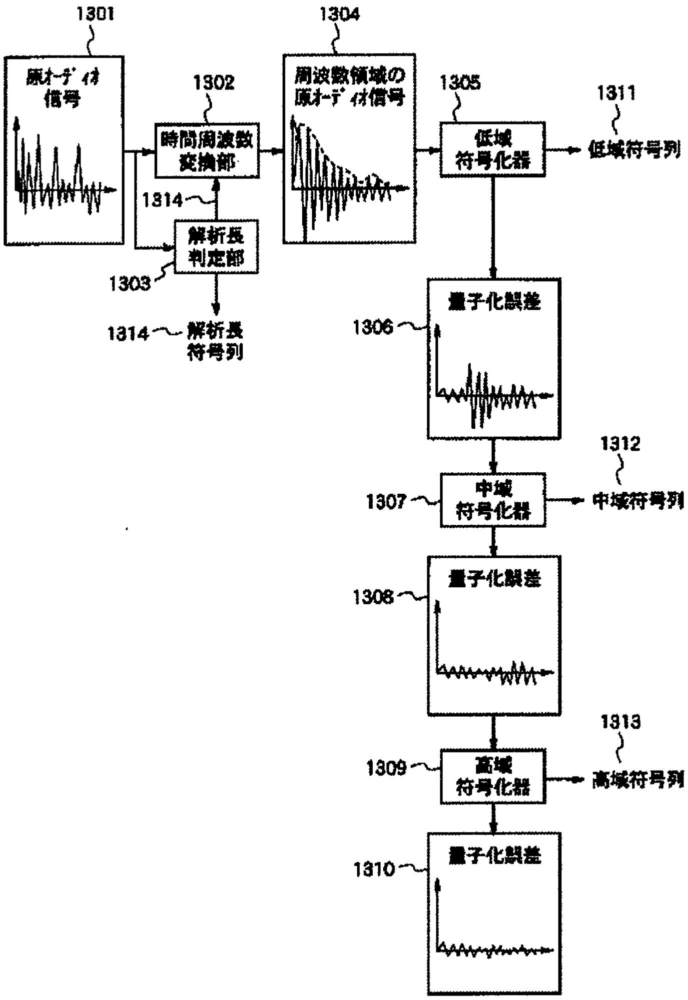
[Drawing 12]



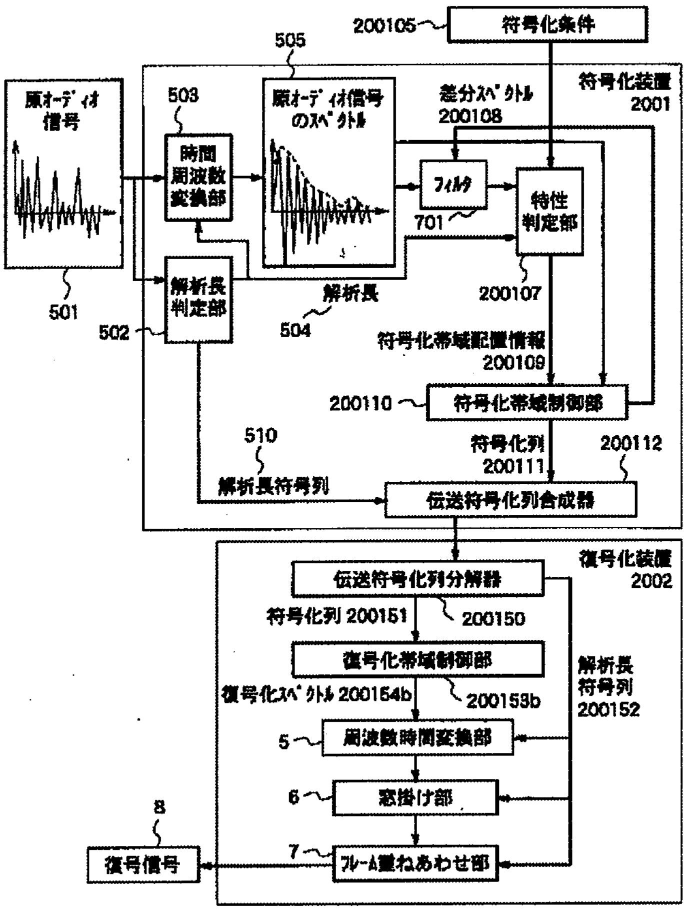
[Drawing 13]



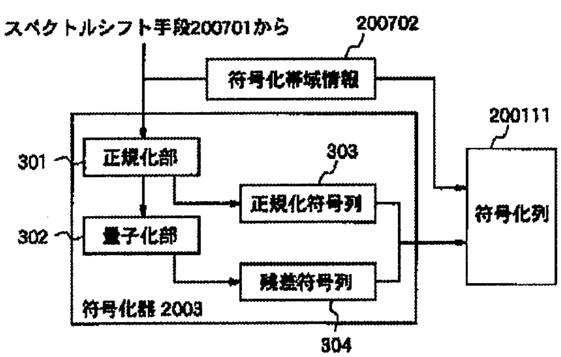
[Drawing 11]


1301
原オーディオ信号
1302
時間周波数変換部
1303
解析長判定部
1314
1314
解析長符号列
1304
周波数領域の原オーディオ信号
1305
低域符号化器
1311
低域符号列
1306
量子化誤差
1307
中域符号化器
1312
中域符号列
1308
量子化誤差
1309
高域符号化器
1313
高域符号列
1310
量子化誤差

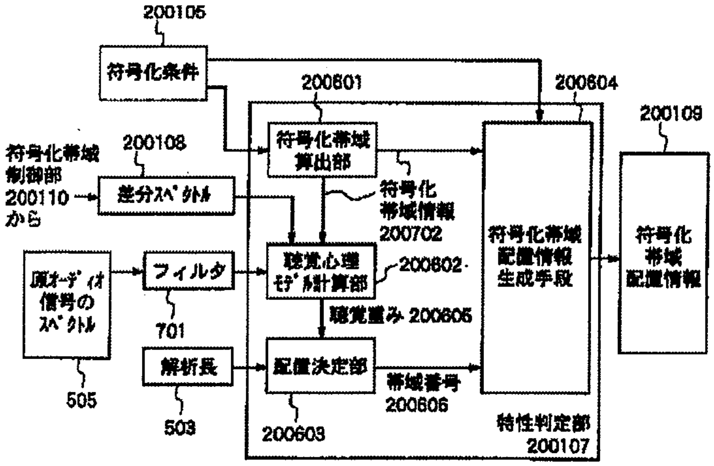
[Drawing 14]



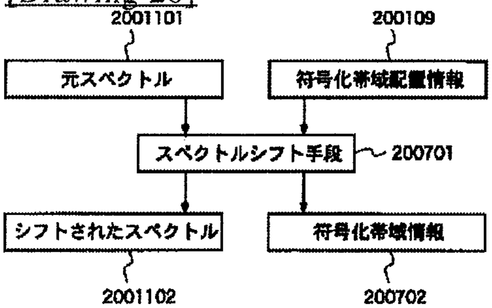
[Drawing 15]



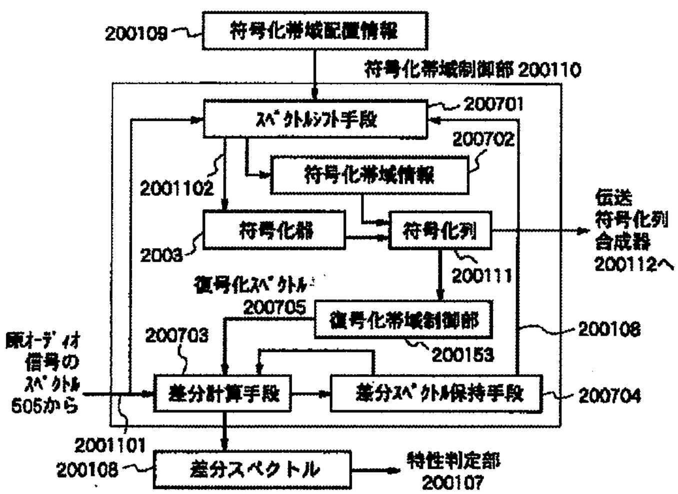
[Drawing 16]



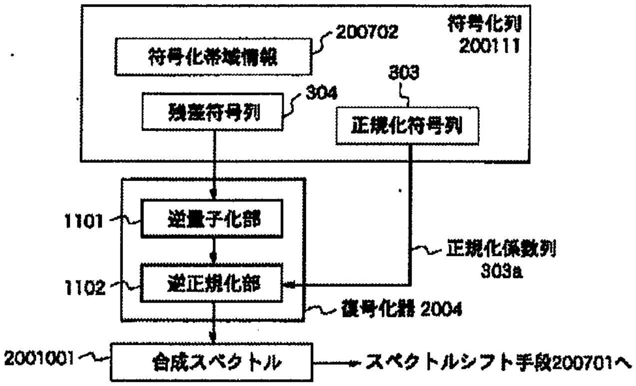
[Drawing 20]



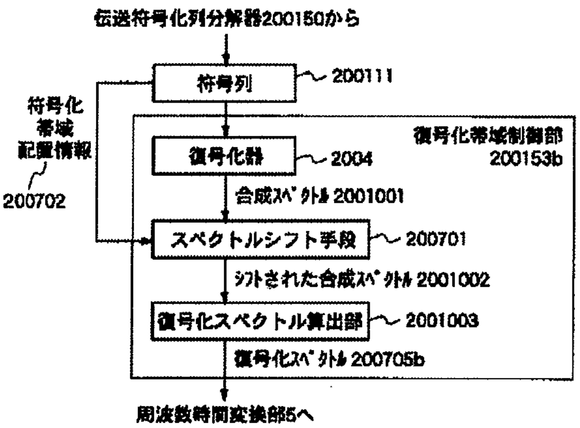
[Drawing 17]



[Drawing 18]



[Drawing 19]



[Translation done.]